AD-A175 942

R-3333-AF

Improving Operational Suitability Through Better Requirements and Testing

William L. Stanley, John L. Birkler

November 1986

A Project AIR FORCE report
prepared for the
United States Air Force

RAND

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER R-3333-AF	2. GOVT ACCESSION NO. A175	3. RECIPIENT'S CATALOG NUMBER 942
4. TITLE (and Subtitle) Improving Operational Suitability Through Better Requirements and Testing		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) W. Stanley, J. Birkler		8. CONTRACT OR GRANT NUMBER(s) F49620-86-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Rand Corporation 1700 Main Street Santa Monica, CA 90406		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Requirements, Programs and Studies Group (AF/RDQM) Ofc, DSC/R&D and Acquisition Hq USAF, Washington, DC 20330		12. REPORT DATE November 1986
		13. NUMBER OF PAGES 104
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

No Restrictions

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Test and Evaluation
Military Requirements
Requirements
Contracts
Acquisition
Weapon Systems
Air Force

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See reverse side

DD FORM 1 JAN 73 1473

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

This report proposes prescriptive actions that could increase the contribution of requirements and test-and-evaluation (T&E) aspects of the weapon system acquisition process to the fielding of more operationally suitable Air Force systems. Actions are needed that (1) correct chronic problems in the expression of operational suitability needs and requirements, (2) address the problem of fragmented operational requirements documentation, (3) expand contractual accountability for reliability-and-maintainability and logistics-support characteristics, (4) adjust acquisition policies to enhance T&E's contribution to decisionmaking and to the identification and correction of deficiencies, and (5) structure tests to demonstrate new operating concepts and capabilities. These actions could facilitate the consideration of suitability factors in acquisition process activities that must address difficult tradeoffs among operational suitability, functional performance, cost, and development time.

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

PREFACE

This report documents the final results of the Project AIR FORCE study "More Supportable Systems Through Requirements/Test and Evaluation Process Improvements," sponsored by the Director of Operations, Office of the Deputy Chief of Staff for Plans and Operations, Headquarters United States Air Force. The overall purpose of the study is to identify acquisition approaches—with emphasis on requirements and test and evaluation (T&E) activities—that could contribute to the development of weapon systems with better operational suitability.

Interim study findings have already been used to help the Air Force improve (a) the expression of specific *Statements of Operational Needs* and (b) general policy guidance and procedures for expressing and documenting operational suitability needs and requirements.

This work is one part of the Project AIR FORCE Resource Management Program's integrated agenda of research aimed at assuring the combat readiness of future forces. The findings and recommendations of the study are intended for use by the many Air Force organizations involved in implementing recent policy guidance by Air Force leadership directing that reliability and maintainability be made primary considerations in the weapon system acquisition process.

SUMMARY

Operational suitability[1] characteristics influence the ability of tactical aircraft to deploy rapidly to combat theaters, to conduct sustained mission operations, to perform to design specifications during those missions, and to be resilient to damage and disruption caused by enemy attacks. Current Air Force systems have suitability characteristics that can detract from their ability to accomplish these tasks. Fielding systems that are more operationally suitable is expected to become more challenging in the future as the increasingly global nature of the enemy threat, and substantial improvements in the size and capabilities of its forces, combine to create appreciable operating environment uncertainties. Developing forces that can operate in these environments will require initiatives that cut across many functional areas in the Air Force, including the requirements and test and evaluation (T&E) processes.

The Air Force can increase the contribution of requirements and T&E processes to the fielding of more operationally suitable systems by (1) correcting chronic problems in the expression of operational suitability needs and requirements, (2) addressing the problem of fragmented operational requirements documentation, (3) expanding contractual accountability for reliability and maintainability (R&M) and logistics support characteristics, (4) considering changes to acquisition policies to enhance T&E's contribution to decisionmaking and to the identification and correction of deficiencies, and (5) structuring tests to demonstrate new operating concepts and capabilities.

The expression and documentation of needs and requirements and the planning and conduct of T&E and the reporting of its results are key acquisition process activities. Statements of requirements provide the basic direction for subsequent acquisition activities. When requirements are poorly stated in the beginning of a program, there is little assurance that the resulting system will meet a user's needs. Similarly, field testing is important because it often provides the first reliable indication of whether a system will in fact meet operational suitability requirements, and of what deficiencies need correction.

---

[1]Department of Defense (DoD) Directive 5000.3 defines operational suitability as "The degree to which a system can be placed satisfactorily in field use, with consideration being given to availability, compatibility, transportability, interoperability, reliability, wartime usage rates, maintainability, safety, human factors, manpower supportability, logistic supportability, and training requirements."

The suggestions made by this report should be regarded as a starting point for enhancing the contribution of requirements and T&E activities to improving operational suitability, because the efficacy of many of the apparently desirable changes in these activities has not yet been fully demonstrated in acquisition programs. Our recommendations stem from reviews of current and past problems with the treatment of operational suitability, assessments of future suitability needs, and our identification of logical responses to address those problems and needs. Programs now in full-scale development, such as the AMRAAM missile and the C-17 aircraft, are providing a test of some the changes identified as potentially desirable by this report, but it will be some time before program outcomes are known.

The nature and extent of operational suitability improvement needed should be decided on a case-by-case basis by examining the criticality of operational suitability to the military capability being sought. When a need for improvement in suitability characteristics is indicated, achieving those improvements will depend not only on effective statements of requirements reinforced with rigorous testing, but also on the willingness of users and developers to face, head-on, difficult tradeoffs among operational suitability, functional performance, cost, and development time. Giving greater visibility to suitability factors in the acquisition process, such as by quantifying more suitability factors in expressions of requirements, can facilitate their consideration in those tradeoff deliberations.

EXPRESSING OPERATIONAL SUITABILITY NEEDS

Historically, Statements of Operational Needs (SONs) have not expressed operational suitability needs well enough to provide guidance to the acquisition and support communities. *Statements of Operational Needs (SONs)* formally document a user's need for a capability to perform military tasks that cannot be satisfied with existing and planned capabilities. Most SONs have focused on deficiencies in functional performance (e.g., speed, maneuverability, range). The treatment of operational suitability needs in SONs could be improved by (1) expanding the aspects of suitability covered, (2) using more operationally meaningful measures that emphasize military capability outputs desired and significant operating constraints, (3) making greater use of quantitative measures of suitability, and (4) prioritizing key needs. Some improvement in the expression of suitability needs is evident in a few recent high-visibility programs, such as the Advanced Tactical Fighter.

EXPRESSING OPERATIONAL SUITABILITY REQUIREMENTS

The expansion and further definition of operational needs as operational requirements that also describe characteristics of the proposed solution to the need could be much improved through the greater use of measures that describe the interactions between R&M and logistics support characteristics that influence the supportability of weapon systems. Operational measures of reliability, for example, should describe not only the frequency at which inherent design failures occur, but also the frequency of occurrence of maintenance activity brought about by maintainability factors such as the fault detection and fault isolation characteristics of equipment. Existing policy guidance identifies and defines many suitability measures, but offers comparatively less assistance in selecting from among those many measures a cohesive set that characterizes the major suitability needs for a particular mission application. Guidance should link appropriate measures to typical usage patterns of major categories of equipment.

DOCUMENTING OPERATIONAL REQUIREMENTS

Improved procedures for formally documenting operational suitability requirements are needed to complement improvements in the substantive expression of requirements. No single document analogous to the *SON* serves as a unified, unambiguous source of Air Force operational requirements. Documentation is currently fragmented across many sources; requirements are inconsistent from one document to another; and it is extremely difficult to correlate key operational, contractual, and test requirements. Institutionalizing use of the *Baseline Correlation Matrix* document, now being applied to selective programs, would begin to address some of the basic documentation problems. Over the long term, development of a unified source of requirements may be desirable that consolidates operational requirements in a single document. A new acquisition document is not necessarily required; with more care in its preparation and review, the *System Operational Concept* document might satisfy the need for better operational requirements documentation.

EXPRESSING CONTRACTUAL REQUIREMENTS

Historically, contractual accountability for suitability-related system characteristics, including R&M, has been quite limited. Contract specifications should more fully reflect the spectrum of maintenance demands that drive the Air Force support burden. They should address a broader set of suitability-related requirements, including aspects of mobility and resilience to attack. Greater use of system-level production equipment specifications appears desirable. The scope and duration of compliance testing should be expanded to support more demanding requirements, and testing should not be compromised by excessive test ground-rule exclusions of potentially important maintenance events. Contractors can be given stronger incentives to meet specifications. The R&M contract specification approach used in the C-17 program displays several promising initiatives in these directions. However, the efficacy of these initiatives will not be fully evident until the 1990s.

USING T&E TO REINFORCE MORE DEMANDING REQUIREMENTS

Acquisition approaches for developing critical subsystems and for managing the transition from development to production limit the contribution of testing. Because of pressures to get systems of higher advertised capability into the field quickly, and desires to avoid apparently costly gaps between development and high-rate production, program schedules have often become too compressed to permit the T&E community to supply significant early input to decisionmakers about a system's operational suitability characteristics and for the development community to make effective use of test information to correct deficiencies before deploying systems in substantial numbers. Highly compressed schedules may permit the testing of characteristics related to functional performance, but many facets of operational suitability (e.g., reliability) require more time to evaluate. The result has often been the fielding of equipment with immature R&M characteristics. Adopting for critical subsystems (such as avionics) an iterative "maturational development" approach that begins before full weapon system development could give the test community systems of greater maturity to test before major acquisition decisions. A "phased acquisition" strategy that extends the low-rate production phase to permit intensive testing and deficiency correction would permit more effective use of knowledge gained from testing to mature systems before beginning production at high rates. Despite the potential of maturational development and phased acquisition for enhancing the contribution of T&E, both face formidable implementation problems involving difficult tradeoffs among system maturity, functional performance, cost, and development time.

More demanding operational suitability requirements will necessitate quantitative and qualitative changes in testing. Future systems will have to possess better reliability and fault isolation characteristics to function effectively in the more stringent operating environments predicted for the future. To demonstrate the achievement of increased reliability will require the accumulation of more test hours, and new kinds of tests and changes in test emphasis will be required to demonstrate new operating concepts and capabilities. For instance, more testing away from main operating base environments will be needed to demonstrate the basing flexibility of future systems and operations in austere environments. Structured field evaluations will be required to demonstrate that systems can sustain high mission-effective sortie rates.

IMPLEMENTATION

The implementation of requirements and testing initiatives is complicated by the diffuse responsibility for operational suitability across the Air Force, which means that solutions require actions by many Air Force organizations rather than by a single office. Air Staff and command organizations recently established as advocates and focal points for R&M (with, hopefully, a broader view about operational suitability as a whole) can serve an important role, however, by encouraging a more disciplined approach to operational suitability. This can help make R&M a primary consideration in weapon system design, as Air Force leadership has directed.

ACKNOWLEDGMENTS

This research has profited from the counsel and assistance of many persons in the Air Force, industry, and RAND. Major General Richard Phillips, while Commander of the Air Force Operational Test and Evaluation Center, and his staff, provided support and encouragement from the inception of the study to its conclusion. Major General William Gorton, while Director of Operational Requirements, Headquarters United States Air Force, encouraged us to apply lessons learned in our research to review the *For Comment Statement of Operational Need for the Advanced Tactical Fighter*. Colonel Harley Garrett, Jr., while heading the Requirements, Programs, and Studies Group in the same directorate, provided useful insights into the requirements process and helped in the acquisition of requirements documents used during early phases of the study. Numerous Air Force personnel in weapon system program offices at the Aeronautical Systems Division and the Armament Division of the Air Force Systems Command shared their insights about requirements and testing.

The contractor community provided a different perspective on requirements and testing issues. Special thanks go to Erv Heald, Shel Hess, and Stu Hann of Douglas Aircraft Company's C-17 program for describing Douglas's approach to satisfying contractual requirements for reliability, maintainability, and availability.

While serving as Director of RAND'S Resource Management Program for Project AIR FORCE, Michael Rich originally suggested this study and provided encouragement and support. Brent Bradley critiqued project briefings. Hyman Shulman and Jean Gebman provided useful comments about the project focus and shared the results of their on-going avionics acquisition studies. Morton Berman served as a technical reviewer of this report. Results from his complementary study on methodologies for determining basing, support, and air vehicle requirements were used in this research. Giles Smith's assessment of the F-16 program contracting approach for another RAND study contributed to this research. Major Alan Ray (while stationed at RAND as an Air Force Research Fellow), Jack Zwanziger, and Edmund Dews assisted in the early phases of the work. Phil Dadant made a thoughtful technical review. Mary Vaiana and Rick Eden helped to sharpen the focus of project briefings and the initial draft of this report.

CONTENTS

FIGURES

TABLES

GLOSSARY

AA	Air-to-Air
ACVP	Administratively Closed, Verification Pending
ACVT	Administratively Closed, T.O. Verification Pending
AFB	Air Force Base
AFFTC	Air Force Flight Test Center
AFOTEC	Air Force Operational Test and Evaluation Center
AFP	Air Force Pamphlet
AFR	Air Force Regulation
AFSARC	Air Force System Acquisition Review Council
AFSC	Air Force Systems Command
AFTEC	Air Force Test and Evaluation Center
AGE	Aerospace Ground Equipment
AIS	Avionics Intermediate Shop
AMRAAM	Advanced Medium Range Air-to-Air Missile
ATARS	Advanced Tactical Aerial Reconnaissance System
ATE	Automatic Test Equipment
ATF	Advanced Tactical Fighter
AWACS	Airborne Warning and Control System
BCM	Baseline Correlation Matrix
BIT	Built-in Test
BIT	Built-in Test Detection
BIT	Built-in Test False Indication
BIT	Built-in Test Isolation
COMO	Combat Oriented Maintenance Organization
CRISP	Computer Resources Integrated Support Plan
DCP	Decision Coordinating Paper
DOC	Designed Operational Capability
DSARC	Defense System Acquisition Review Council
DT&E	Development Test and Evaluation
Dev	Development
DoD	Department of Defense
DoDD	Department of Defense Directive
ECP	Engineering Change Proposal
FAC	Forward Air Control
FD	Fault Detection
FI	Fault Isolation
FMC	Full Mission Capable
FOD	Foreign Object Damage
FOL	Forward Operating Location

FOT&E	Follow-on Operational Test and Evaluation
FSD	Full-Scale Development
GSE	Ground Support Equipment
ICS	Interim Contractor Support
IFF	Identification Friend or Foe
ILSP	Integrated Logistics Support Plan
IOC	Initial Operational Capability
IOD	Initial Operational Delivery
IOT&E	Initial Operational Test and Evaluation
IPS	Integrated Program Summary
JMSNS	Justification for Major System New Start
JSOR	Joint Service Operational Requirement
LRU	Line Replaceable Unit
LTE	Launch to Eject
MAJCOM	Major Command
MC	Mission Capable
MCSP	Mission Completion Success Probability
MDT/S	Mean Downtime per Sortie
MER	Multiple Ejection Rack
MIL	Military Standard
MMAX	Maximum Time to Repair
MMH/FH	Maintenance Man-hours per Flying Hour
MMH/FH(C)	Maintenance Man-hours per Flying Hour (Corrective)
MMTR	Mean Man-hours to Repair
MOB	Main Operating Base
MOE	Measure of Effectiveness
MOT&E	Multinational Operational Test and Evaluation
MRT	Mean Repair Time
MTBCF	Mean Time Between Critical Failures
MTBD	Mean Time Between Demands
MTBF	Mean Time Between Failures
MTBM	Mean Time Between Maintenance
MTBM(C)	Mean Time Between Maintenance (Corrective)
MTBM(I)	Mean Time Between Maintenance (Inherent)
MTBMA	Mean Time Between Maintenance Actions
MTBR	Mean Time Between Removals
MTTF	Mean Time to Failure
MTTR	Mean Time to Repair
MTTRS	Mean Time to Return to Service
NATO	North Atlantic Treaty Organization
NAVMAT	Naval Material
NMC	Not Mission Capable
NMCB	Not Mission Capable, Both (Maintenance and Supply)

NMCM	Not Mission Capable, Maintenance
NMCS	Not Mission Capable, Supply
OHMS	On-board Health Monitoring System
OSD	Office, Secretary of Defense
OT&E	Operational Test and Evaluation
Opnl	Operational
PAA	Primary Aircraft Authorized
PAR	Program Assessment Review
PDR	Preliminary Design Review
PMC	Partially Mission Capable
PMCB	Partially Mission Capable, Both (Maintenance and Supply)
PMCM	Partially Mission Capable, Maintenance
PMCS	Partially Mission Capable, Supply
PMD	Program Management Directive
PMP	Program Management Plan
POL	Petroleum, Oil, Lubricants
PSOC	Preliminary System Operational Concept
QTA	Quick Turnaround
R&D	Research and Development
R&M	Reliability and Maintainability
RADC	Rome Air Development Center
RFP	Request for Proposal
ROC	Required Operational Capability
RTOK	Re-test OK
SAR	Selected Acquisition Report
SCP	System Concept Paper
SDDM	Secretary of Defense Decision Memorandum
SEDS	System Effectiveness Data System
SOC	System Operational Concept
SON	Statement of Operational Need
SOW	Statement of Work
SPO	System Program Office
SPR	System Program Review
Sp	Special Operations
T&E	Test and Evaluation
TAC	Tactical Air Command
TCTO	Time Compliance Technical Order
TD	Test Directive
TEMP	Test and Evaluation Master Plan
TER	Triple Ejection Rack
TEWS	Tactical Electronic Warfare System
TO	Technical Order

UHF	Ultra High Frequency
U.S.	United States
USAF	United States Air Force
UTC	Unit Type Code
VHF	Very High Frequency
VTOL	Vertical Takeoff and Landing
WRSK	War Reserve Spares Kit
WSR	Weapon System Reliability
WUC	Work Unit Code

I. INTRODUCTION

BACKGROUND

The need to remedy performance deficiencies (in speed, maneuverability, lethality, etc.) in the combat arena has been, and probably always will be, the principal reason for developing new weapon systems, but the defense community is increasingly recognizing that the multidimensional set of characteristics commonly grouped under the heading of operational suitability[1] is also an important contributor to overall mission accomplishment of weapon systems in wartime. Reliability, maintainability, logistics and manpower supportability, and a host of other operational suitability characteristics influence the ability of tactical weapon systems and their associated support to rapidly deploy to combat theaters, to conduct mission operations on a sustained basis, to perform to design specifications during those missions, and to be resilient to damage and disruption caused by enemy attacks.

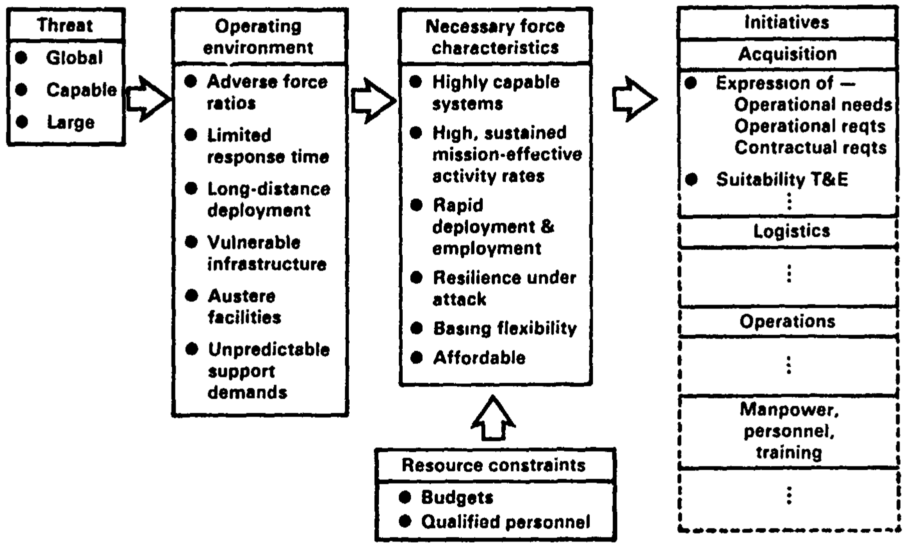
Some contemporary Air Force weapon systems have operational suitability characteristics that can make wartime support in the field difficult and detract from the full realization of designed combat performance. During peacetime, these same characteristics can increase operating and support costs, which in turn can limit the number of systems that can ultimately be bought and/or the quality of their support, and can make it harder to accomplish the peacetime training mission.[2]

Fielding more operationally suitable systems is expected to become more challenging in the future as the increasingly global nature of the enemy threat and substantial improvements in the size and capabilities of its forces combine to create appreciable operating environment uncertainties (see Fig. 1). Sizable U.S. forces may have to deploy rapidly over long distances with little warning and operate in locations having only limited support facilities, such as Third World locations, dispersed operating locations in the European theater or elsewhere, or

---

[1]Department of Defense (DoD) Directive 5000.3 defines operational suitability as "The degree to which a system can be placed satisfactorily in field use, with consideration being given to availability, compatibility, transportability, interoperability, reliability, wartime usage rates, maintainability, safety, human factors, manpower supportability, logistic supportability, and training requirements."

[2]For more details about the operational suitability characteristics of contemporary equipment, see Rich, Stanley, and Anderson, 1984.



SOURCE: Rich et al. (1984).

Fig. 1—Factors driving the need for more supportable systems

otherwise well-equipped operating bases in NATO whose support facilities have been damaged or disrupted by enemy attack.[3]

These challenging operating environments, together with continuing budget constraints and a shrinking pool of qualified personnel, will shape the kinds of capabilities and operating characteristics needed for future forces. U.S. forces will continue to need highly capable weapon systems that depend on complex subsystems to counter enemy advances in quality and numbers. Rapid and sustained fully-mission-effective operations will be needed to cope with adverse force ratios. This capability should be resilient to damage and disruption at our bases and facilities and be achievable with a minimum of deployed resources so as to allow for sudden deployments, permit operations in austere environments and dispersed modes, and facilitate reconstitu-

[3]See Rich et al. (1984) for a more comprehensive treatment of the enemy threat facing the Tactical Air Forces in the next several decades, how that threat affects the combat environments within which those forces will have to operate, and how those environments should shape force characteristics.

tion after being stressed. Moreover, these force characteristics must be affordable.[4]

Achieving these force characteristics will require an integrated set of actions by the Air Force that cut across many functional areas, including the acquisition process, logistics, training, and operations.[5] The adequacy with which each activity in the acquisition process treats operational suitability will have to be examined, from the earliest formulation of requirements through field test and evaluation, that often provides the first reliable indication of whether a system will in fact meet operational suitability requirements, and if not, of what deficiencies need correction.

A September 1984 joint action memorandum from the Secretary of the Air Force and the Chief of Staff of the Air Force established the priority of reliability and maintainability for weapon systems.

> For too long, the reliability and maintainability of our weapon systems have been secondary considerations in the acquisition process. It is time to change this practice and make reliability and maintainability primary considerations.
>
> We must emphasize reliability and maintainability throughout the acquisition process—from requirements definition, through concept development, design, production, and acceptance. Everyone must insure reliability and maintainability requirements are met through every step of the process. Reliability and maintainability must be coequal with cost, schedule, and performance as we bring a system into the Air Force inventory.[6]

Implementing this guidance will, in some cases, involve making some difficult tradeoffs among cost, development time, functional performance, and operational suitability.[7] A willingness on the part of Air Force operators (the ultimate users of systems) and developers to address such tradeoffs head-on is a necessary condition for achieving better operational suitability. Indeed, absence of such a willingness will compromise many other proposals for improving operational suitability, including those discussed in this report.

Improving the expression of operational suitability requirements is one means for encouraging and facilitating the examination of tradeoffs, and is a major focus of this report. Improving suitability requirements begins with the first broad expression of a user's need for a capability to accomplish a particular task derived from mission area

---

[4]Ibid.

[5]Ibid. The report gives a broader perspective of the necessary actions.

[6]Excerpt from Gabriel and Orr (1984).

[7]Methodologies for addressing such tradeoffs are becoming increasingly available. See

analysis, and extends to more definitive expressions in contractual specifications of the characteristics of the solution to the need being sought by the Air Force.

Initial expressions of those needs must strike a delicate balance, giving enough guidance to the research and development (R&D) community to shape design studies, without being so specific and narrowly drawn as to stifle a full examination of potential solutions to the operational need. Decisions about what operational suitability needs to quantify, and what measures to use to accomplish that quantification, become very important. As information becomes available from design studies about the feasibility of satisfying the user's needs and the most attractive means for doing so, acquisition objectives should be adjusted to reflect the new information.

OBJECTIVES

This research sought to identify how improvements in the expression of requirements and in test and evaluation (T&E) activities can contribute to Air Force efforts to improve the operational suitability of its weapon systems. The research had three primary objectives:

- Identify approaches for improving the expression and use of operational suitability needs and requirements.
- Identify ways to make contractual reliability and maintainability requirements statements, and associated compliance measurements, more meaningful.[8]
- Identify changes in the approach to test and evaluation that can reinforce more demanding operational suitability requirements.

Figure 2 shows the interactive relationship of the activities addressed by the objectives. The separation of operational needs and requirements depicted in Fig. 2 belies the fact that there is no universally accepted distinction in the Air Force between needs and requirements. Expressions of operational needs are supposed to describe a user's need for a capability to perform military tasks that cannot be satisfied with his existing and planned capabilities. Ideally, need statements should focus on describing the end result desired by the user, subject to any particular constraints the user must operate under.

for example, Berman (1985).

[8]"More meaningful" refers to making requirements expressions and test measurements bear a closer relationship to the actual operational tasks facing the system in the field. Suggestions for doing so follow in subsequent sections of the report.

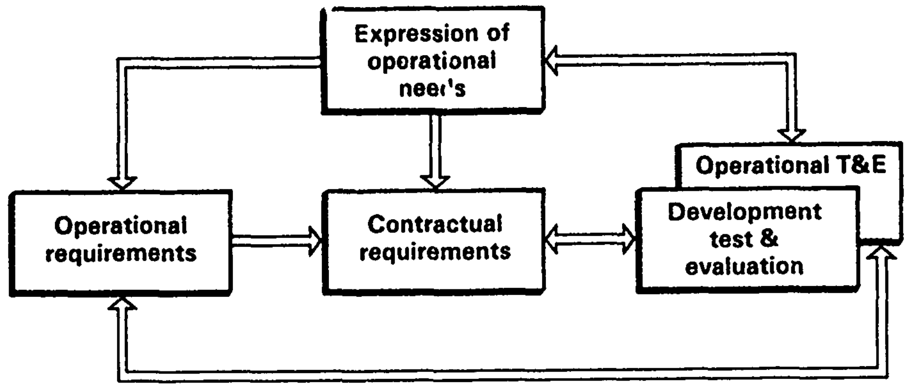


Fig. 2—Relationship of needs, requirements, and testing

Operational requirements statements provide more details about the characteristics of the solution to the user's needs. To do so, those statements may refine, extend, or expand expressions of operational needs. Thus, they provide more details about the means to the end result desired. As should be apparent from these distinctions, there is a considerable gray area involving which kinds of parameters belong in need statements and which belong in statements of operational requirements.[9]

Expressions of operational needs help in the development of more detailed operational requirements, contribute to the formulation of contractual requirements, and are used by the test community for planning and evaluation. System Program Offices (SPOs) use operational requirements statements to develop contractual requirements stating what the weapon system contractor is expected to deliver. The development and operational test communities use requirements for test planning and evaluation. Lessons learned during testing feed back to the expression of needs and requirements.

[9]Section II will describe an approach for deciding which parameters to specify in need statements. Section III will present a framework for selecting measures to describe operational requirements involving reliability, maintainability, and logistics support characteristics.

SCOPE

This research did not examine all requirements and T&E activities that might conceivably influence the many dimensions of operational suitability. Instead, it focused on activities that have the potential for beneficially influencing those system capabilities deemed particularly important for operations in more stringent environments expected in the future, including mission-effective sortie generation (or the corresponding function for non-aircraft systems), basing flexibility, mobility, and resilience to attack.

The research did not assess the cost impacts, pro or con, of changes in requirements and T&E activities, but it did note when particular recommendations could have a cost impact.

APPROACH

To analyze past and present approaches for expressing operational suitability requirements and testing for their achievement, the research used document reviews, discussions with key personnel from many organizations, and coordination with related RAND weapon system acquisition and support research efforts. We reviewed policy guidance that addressed many aspects of operational suitability requirements and testing, including guidance published by the Department of Defense, the Joint Chiefs of Staff, the Navy, Air Force headquarters, major operating commands, and separate operating agencies. We also reviewed numerous program documents for tactical and strategic aircraft and missiles, ranging from initial *Statements of Operational Needs* to contractual system specifications to developmental and operational T&E reports.

We discussed requirements and testing with major operating commands that initially express needs and requirements, SPOs that write contractual requirements and manage the development of systems to meet the requirements, contractors that design to those requirements, test organizations that measure the adequacy of systems, commands responsible for supporting systems in the field, and DoD and Air Force organizations that review program progress. The research effort was also supported by related RAND research in the areas of avionics acquisition and support, airbase modeling, methodologies for determining basing, support, and air vehicle requirements for aircraft weapon systems, and resource management strategies for improving readiness and sustainability.

This research approach helped in identifying desirable improvements in the statement of requirements and in T&E activities. However, it should be acknowledged that many of the apparently desirable changes in approach to requirements and T&E activities that this report recommends, while logical and based on sound reasoning, have not yet been fully demonstrated in the crucible of actual acquisition programs. This reflects the fact that operational suitability considerations have often not occupied a prominent position in program planning in the past. Programs now in full-scale development (e.g., the C-17 transport aircraft, AMRAAM missile) will provide a test of some approaches for improving operational suitability discussed in this report, but it will be some time before we will be able to assess program outcomes. Moreover, because acquisition outcomes are the product of many factors, it will always be difficult to conclusively establish cause and effect relationships. Nonetheless, subject to adjustment as experience dictates, the recommendations developed here provide a starting point for enhancing the contribution of requirements and T&E activities to improving operational suitability.

ORGANIZATION OF REPORT

The organization of this report follows the framework of Fig. 2. **Section II** reviews the historical record of Air Force expression of operational suitability needs, suggests ways to express those needs more effectively, discusses the importance of prioritizing needs, and examines existing policy guidance for expressing suitability needs.

Section III identifies some shortcomings of past statements of operational requirements, illustrates how they can create difficulties for various organizations in the acquisition process, describes a philosophy for the selection of measures to express operational suitability requirements more effectively, and identifies some problems and potential solutions associated with operational requirements documentation.

Section IV identifies some limitations of past contractual statements and associated compliance measurements of support requirements and some potentially promising new approaches, and assesses some of the implications for the contractor and the Air Force of new contracting approaches.

Section V illustrates how existing acquisition approaches limit opportunities for enhancing suitability testing, suggests how acquisition strategy changes can facilitate more meaningful testing and the more effective use of T&E results, and identifies quantitative and qualitative changes in suitability testing needed to demonstrate new operating concepts and capabilities.

Section VI summarizes the research findings and recommendations.

II. EXPRESSING OPERATIONAL SUITABILITY NEEDS

The identification, documentation, and submittal of operational needs is an integral part of the weapon system acquisition process. New and improved weapon systems are the product of research, development, and acquisition programs defined and undertaken to satisfy needs first expressed early in the acquisition process. Thus, the effectiveness with which needs are expressed can have a direct bearing on the characteristics of weapon systems introduced in the field.

This section suggests ways to express operational suitability needs more effectively in Air Force *Statements of Operational Needs (SONs)*. We identify shortcomings in past expressions, illustrate a framework for identifying those needs that appear desirable to quantify, identify measures to express needs, illustrate the importance of prioritizing needs, and critique existing policy guidance for expressing suitability needs.

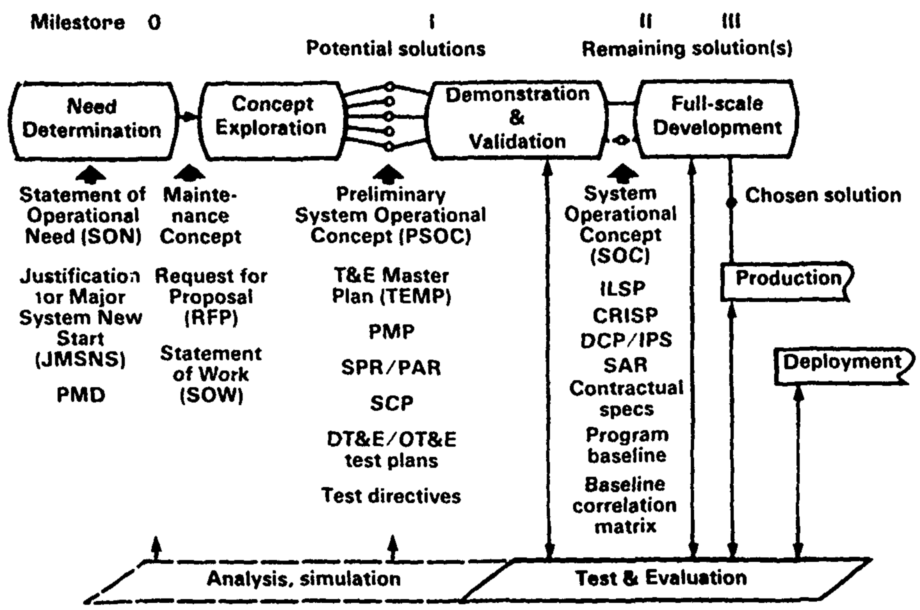
THE IMPORTANCE OF THE SON IN THE ACQUISITION PROCESS

Air Force organizations responsible for particular mission areas identify operational needs through continuing mission area analysis. These needs normally result from deficiencies in the ability to perform the assigned mission or from opportunities to perform the mission better because of technological advances or other factors. The *SON* documents operational needs that Air Force organizations cannot satisfy within their own authority and resources. As one of the first major documents in the genesis of new weapon system concepts, the *SON* sets the tone for subsequent development efforts by the operational deficiencies it emphasizes, the particular needs for new or improved capabilities it identifies, and the way it expresses those needs.

SONs that fail to adequately address or express operational needs can have far-reaching effects that extend beyond the requirements phase. Many acquisition activities involving program justification, design studies, the planning and conduct of test and evaluation, program management, and acquisition decisionmaking use *SONs* either directly or indirectly for guidance (see Fig. 3).[1] Since many of these activities occur before the start of full-scale development, the effective expression of needs early in the acquisition cycle is important.

In supporting the *definition and validation of operational needs*, the *SON* helps the Air Force develop a corporate perspective of its needs and priorities. This process gives visibility to needs, and the competition among various needs for validation and funding in the budget process forces the Air Force to set priorities.[2]

Effective expression of suitability needs is also particularly important to *support early design studies* that make the tradeoffs that effectively shape the suitability characteristics the system will possess. *Requests for Proposals (RFPs)* issued prior to concept exploration draw on the *SON*, and ultimately on operational concept documents, to *give*



SOURCE: Documents as noted.

Fig. 3—The SON in the acquisition process

[1]See the Glossary for meanings of the acronyms in Fig. 3.

[2]Validation of a *SON* implies that the need has been judged sound by the Requirements Review Group at Air Force Headquarters and that the proposed program for satisfying that need will be permitted to compete for funding in the budget process.

guidance to contractors.[3] However, when *SONs* express needs in qualitative terms such as "highly reliable subsystems are required," operational suitability may be traded away in favor of more quantitatively expressed functional performance needs.

SONs support the planning and conduct of test and evaluation through their use as a source of inputs for preparing the *Test and Evaluation Master Plan (TEMP)* and subsequent more detailed test plans that describe the structure of the testing that provides a substantive basis for evaluating systems. Poor expressions of needs can complicate the planning and conduct of T&E and the use of its products for decisionmaking.

Program managers and acquisition decisionmakers use needs and more detailed expressions of requirements in the *management and evaluation of weapon system acquisition programs.* Numerous documents and briefings incorporate information from *SONs.* To the extent that they reflect suitability needs that are well-expressed in *SONs*, these documents encourage, but do not guarantee, management attention to suitability.

How well have past *SONs* supported this community of users? Somewhat surprisingly, given the importance of the document and of operational suitability to mission accomplishment, the historical record of the expression of suitability needs has not been particularly good. However, some improvement is evident in *SONs* for recent high-visibility programs.

HISTORICAL RECORD OF SUITABILITY TREATMENT IN SONs

We reviewed a 19-year sample of need statements[4] for tactical and

---

[3]*RFPs* are generally issued prior to each major program phase. For clarity, Fig. 3 shows only one occurrence. Similarly, other documents, e.g., *TEMP, PMP, DCP*, are periodically updated, but the figure illustrates only their first occurrence.

[4]The sample includes 6 *SONs*, 8 *Required Operational Capability (ROC)* documents that served a similar function in the past, and one *Joint Service Operational Requirement (JSOR)*. Eight of the 15 need-statements have resulted in operational systems that are now in the field; an additional three systems in the sample are in various stages of development. Eleven of the 15 need-statements are for systems that have primarily a tactical orientation; the remainder have a strategic orientation. We have also reviewed the expression of suitability needs for some non-aircraft systems in other *ROCs, SONs*, and *JSORs*. With one notable exception, the need-statements were reviewed solely for research purposes to identify problem areas and to stimulate thinking about how need-statements could be improved. The exception was the *For Comment Statement of Operational Need for the Advanced Tactical Fighter/Air-to-Air (ATF/AA)*, 6 October 1983, which RAND reviewed and commented on at the request of the Director of Operational Requirements, Headquarters United States Air Force. The Tactical Air Forces have subsequently issued a revised *For Coordination* version of this *SON*, 18 October 1984, to which the comments in this report refer.

strategic aircraft (see Table 1). They were evaluated in terms of their

- Acknowledgment of suitability as one basis of the need,
- Identification of operational suitability deficiencies in current systems,
- Identification of technological opportunities for enhancing suitability,
- Coverage of key suitability needs and their quantification,
- Completeness of operational setting descriptions,
- Adequacy of measures selected to express suitability needs, and
- The extent of prioritization.

Table 1

AIRCRAFT NEED-STATEMENTS REVIEWED

Need	Document Type	Year Issued	System	Status
AWACS	ROC	1966	E-3	Opnl
Advanced Manned Strategic Aircraft	ROC	1966/78	B-1	Opnl
Advanced Tactical Fighter (F-X)	ROC	1968	F-15	Opnl
Wild Weasel	ROC	1968	F-4G	Opnl
EF-111A TEWS	ROC	1973	EF-111A	Opnl
Advanced Multipurpose Tanker	ROC	1973	KC-10	Opnl
Improved FAC Aircraft (FAC-X)	ROC	1976		Pre-dev
F-16 Aircraft	ROC	1976	F-16	Opnl
Cruise Missile Carrier Aircraft	SON	1979		Cancelled?
Near-Term Manned Bomber	SON	1979	FB-111H	Cancelled
Special Operations VTOL (H-X)	SON	1979	HH-60D	Dev
Continuous Battlefield Surveillance	SON	1979	TR-1	Opnl
Advanced Tactical Aerial Reconnaissance System	SON	1979		Pre-dev
JVX	JSOR	1982	V-22	Dev
Advanced Tactical Fighter (Air-to-Air)	SON (For Coord)	1984		Dev

SOURCE: ROCs, SONs, JSORs as noted.

In all cases, the focus was on evaluating how needs were expressed so we could identify problem areas, and not on critiquing the correctness of the need itself nor any numerical value associated with its expression.

Citing Better Suitability As a Basis For a Need

The need for better operational suitability characteristics has rarely been cited as an important justification for developing a new capability—the *ATF SON* is an exception. Moreover, most need statements have not addressed how the threat shapes desired operational suitability characteristics. The growth in enemy attack capabilities is likely to make the threat of damage and disruption to aircraft support resources an important consideration in specifying needed operational suitability characteristics in future *SONs*.[5]

Identifying Suitability Deficiencies in Current Systems

Most need-statements have focused almost exclusively on shortfalls in the functional performance of existing systems, such as speed, maneuverability, range, or lethality. By not explicitly identifying operational suitability deficiencies in existing systems, these need-statements missed opportunities to stimulate the investigation of design options that might overcome those deficiencies with the new system or capability being requested, and could perhaps be interpreted as implying satisfaction with the suitability characteristics of existing systems. The identification of deficiencies is particularly important because the design community does not always have a good appreciation of problems brought about by undesirable operational suitability characteristics and experienced by users in the field. Users, with support from the logistics community, can supply this important input.

The need-statement for the Advanced Tactical Fighter (ATF) is exceptional in that it does discuss the operational suitability deficiencies of current forces that need to be addressed in designing the ATF. It identifies as deficiencies such things as large support requirements when deployed, large specialized workforces, hard-to-handle materials, high failure rates, fault isolation problems, etc., that in various ways detract from mobility, sortie generation, and resilience to attack.

[5]See, for example, Rich et al. (1984) and Emerson (1982).

Identifying Technological Opportunities for Enhancing Suitability

Need statements have usually not provided a balanced discussion of technological opportunities that can enhance both effectiveness *and* suitability. Such discussion can serve as one impetus for the design community to evaluate those technologies. Discussion usually focuses on those technological advances that may contribute to better effectiveness in the combat arena—for example, engines with higher thrust-to-weight ratios, thrust reversing and vectoring, advanced flight control concepts, aerodynamically efficient wings, lighter composite structures, advanced weapons, etc. The *ATF SON* does acknowledge that fault-and-damage-tolerant avionics architectures could improve reliability and permit the use of smaller, less complex support equipment. None of the documents we reviewed encouraged evaluation of the potential of technology and design innovations for incorporating built-in support—a function now supplied by extensive equipment on flight lines.

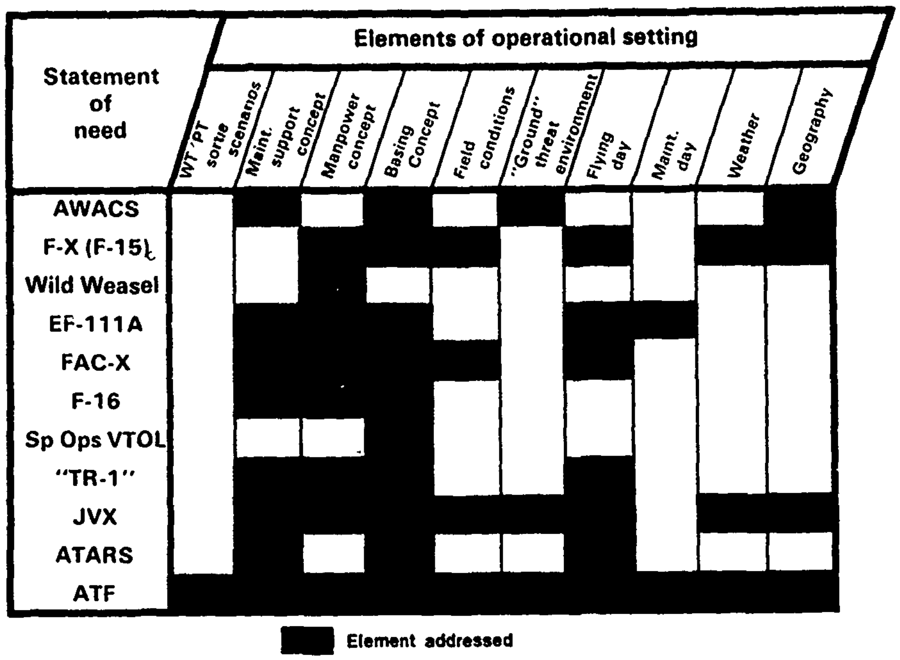
Coverage and Quantification of Suitability Needs

We believe that effective expressions of operational suitability needs should:

- Be accompanied by a contextually complete description of the operating environment under which the measure applies;
- Measure operationally meaningful characteristics;
- Be as quantitative and/or specific as possible;
- Be linked to contractual terminology; and
- Be measurable through combinations of analysis, evaluation, simulation, and field testing

We reviewed the eleven tactical aircraft need-statements to determine whether they described key aspects of the operating environment.[6] Some aspects of the environment have rarely been specified, and the coarse characterization of Fig. 4 masks considerable variability even when a particular aspect of the environment is addressed. For example, with regard to field conditions, the F-X document merely stated a takeoff and landing distance requirement, from which one might infer a field length, whereas the ATF document specifies the length of runways from which the aircraft must operate, the field altitude, weather conditions, and the bearing strength of the surface.

[6] "Described" in this context means at least mentioned or discussed, although not necessarily quantified. Not all aspects of the operational setting description are amenable to quantification.



SOURCE: ROCs and SONs as noted.

Fig. 4—Operating environment descriptions in SONs

Generally, the *ATF SON* specifies the operating environment better than any of the other aircraft need-statements.[7]

It is a thesis of this research that quantification of operational suitability needs is desirable and necessary to get the R&D community to respond in a meaningful way to the call issued by Air Force leadership for better R&M (and, by extension, operational suitability), to focus development attention on key aspects of operational suitability by providing a substantive *starting point* for design and tradeoff studies, and to provide a framework for evaluating the attractiveness of alternative concepts for satisfying needs. While not necessarily condoning a lack of quantification in *SONs*, those responsible for writing, reviewing, and

[7]Policy guidance for the *System Operational Concept (SOC)* document has historically contained rather detailed checklists for expressing many aspects of the operational setting. Although the level of detail called for in the *SOC* is more than that which would normally be incorporated in a *SON*, such guidance can serve as a useful reference. At this writing, policy guidance for the *SOC* is being split away from guidance for the *SON* and will be forthcoming in AFR 55-24, *Concepts, System Operational Concepts.*

using need-statements to whom we spoke cited a number of reasons why they believe there is an apparent reluctance to quantify suitability needs. Such arguments provide a context both for our historical review of the quantification of needs in *SONs* that immediately follows and subsequent recommendations about how to identify needs to quantify and the measures to use to accomplish that quantification.

First, there is concern that quantifying suitability needs could constrain designers in such a way as to preclude their exploration of a full range of potential design solutions. Second, some assert that there is simply not enough information early in the acquisition cycle about the implications of satisfying particular suitability needs to make informed estimates of the values of key operational suitability characteristics. Third, it has been suggested that the confusing terminology used to describe various R&M characteristics discourages efforts to quantify suitability needs. Fourth, there is a concern that quantified needs can take on a life of their own, independent of what subsequent design studies may indicate, creating the potential for programmatic difficulties later in acquisition during testing and program reviews.

Other reasons for a reluctance to quantify suitability needs are less openly acknowledged. There is concern that satisfying an explicitly quantified suitability need may result in some functional performance being traded away. There is also concern that satisfying a quantified suitability need may drive up costs to the point that the program for satisfying the need is endangered. Finally, it is argued that there is more operating flexibility for subsequently changing needs when they are expressed qualitatively.

We believe the aforementioned concerns ignore some realities about how *SONs* typically evolve and ignore some alternative approaches for quantifying needs that can mitigate some of the perceived adverse consequences from quantifying suitability needs.

First, an approach that emphasizes the quantification of needs representing the end result desired by the user rather than the means to the end (i.e., the characteristics of the hardware solution to the need) can provide sufficient flexibility to the development community to examine a sprectrum of design solutions. We will illustrate this subsequently when we discuss an approach for quantifying tactical aircraft suitability needs.

Second, it is important to recognize that the quantification of operational suitability needs is not (or need not be) accomplished in a vacuum. In anticipation of the development of operational needs, development planning organizations usually conduct or sponsor preliminary trade studies to develop some general notions about what are feasible solutions for satisfying anticipated needs. In any event, contractors are conducting on-going studies to prepare for design competitions, and these results are frequently communicated to the user community. Development organizations also are performing comparability analyses to help in defining reasonable expectations about R&M performance of future systems. Air Force laboratories also make technology assessments to help in defining what may be feasible. Finally, users are conducting mission area analyses to define the need for new or improved capabilities.

Third, the evolution of *SONs* is an iterative process which affords a broad cross-section of organizations the opportunity to critique needs quantified by the user. While a *SON* is seldom changed after a need is finally approved, the document goes through several stages that can extend over several years or more.[8] Initially, a *SON* is circulated in a "For Comment" version while concept development studies are under way prior to the DSARC or AFSARC I decision milestone, at which time a system's readiness for the demonstration and evaluation phase (advanced development) is evaluated. The needs stated in that document are critiqued by Air Force organizations and either directly or indirectly by contractors as well. In particular, the development community comments on the technical feasibility of satisfying the need. Another "For Comment" version may then be issued after the user has revised the *SON* in accordance with the comments. Ultimately, a "For Coordination" version is issued for signature approval by involved organizations, but only after considerable opportunity for feedback on what the user is asking for. This is not to say that the review process precludes unreasonable quantified needs from surviving in approved *SONs*, but at least there is a structure that affords a large community the opportunity to comment.

Fourth, because the *SON* evolves in an iterative fashion, many of the quantified needs in it should properly be regarded as goals or objectives, unless there is a compelling reason that requires that a quantified need be regarded as an absolute threshold of minimum acceptable performance. Ideally, in the "For Comment" version of *SONs*, quantified suitability needs should act as benchmarks about which design tradeoff studies are accomplished to explore the potential implications of those needs on other aspects of system performance, cost, and schedule. The exception may be those needs driven by threat factors or intractable operating constraints that are apparent even early in the evolution of the need.

---

[8]For a complete description of the evolution of the *SON*, see *Operational Requirements, Operational Needs*, Air Force Regulation 57-1.

Fifth, the *SON* is not the last word on the need that shapes the design solution. Even after a *SON* is approved, knowledge gained subsequent to its approval may require some adjustment in acquisition objectives that can be reflected in subsequent requirements documents, such as the *System Operational Concept (SOC)*, which is updated as necessary before each decision milestone.

On balance, we believe that users should be mindful of potential pitfalls in quantifying suitability needs, but that by carefully drafting need-statements and by subjecting them to thorough review by the acquisition community, they can generate much more meaningful expressions of suitability needs. Such an approach will represent a departure from precedent, for our review of a 19-year sample of tactical and strategic aircraft need-statements illustrates that many operational suitability characteristics have been infrequently or inconsistently quantified, although we cannot estimate whether this is more a consequence of deliberate action because of the concerns about quantification noted above, or because of simple neglect. Quantification is inconsistent or infrequent whether one considers quantification of top-level measures such as sortie generation or availability, or intermediate measures such as system reliability and maintainability that influence a system's availability and ability to generate sorties (see Fig. 5). At times, solution-oriented characteristics are quantified (e.g., system reliability and maintainability), while important operating constraints such as manpower levels and limits on deployed support, which influence mobility, are not mentioned. Need-statements for the four strategic systems at least quantified availability needs, a particularly important attribute for such systems, but only one of the eleven tactical need-statements, the *ATF SON*, quantified sortie generation needs. No consistent strategy has been applied to identify needs requiring quantification.

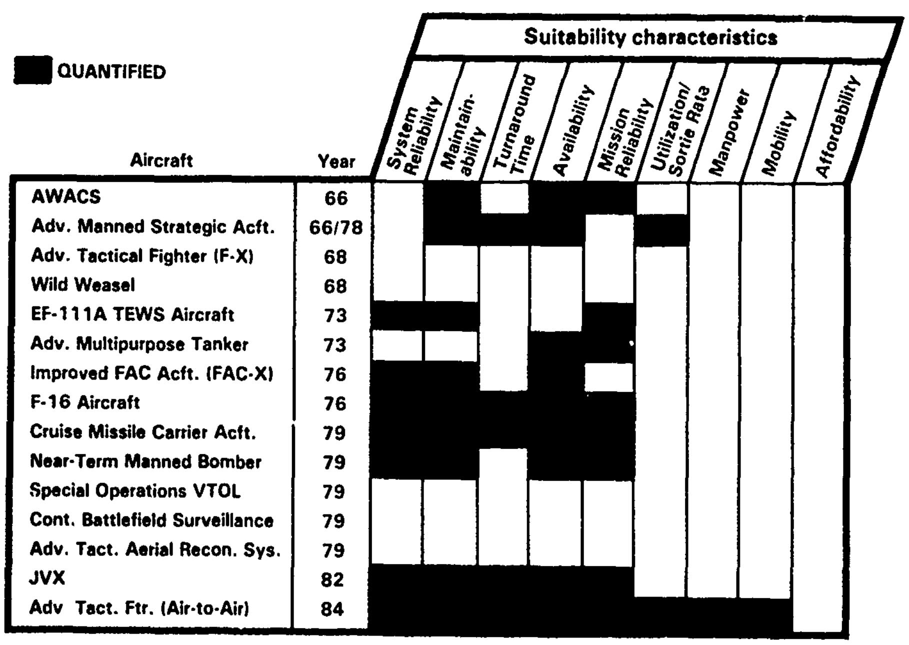
The *ROC* that led to the F-15 is symptomatic of the comparative neglect in the expression of suitability needs relative to the expression of functional performance needs in past statements. It contained a single qualitative paragraph about suitability needs that possessed none of the traits noted above. It did, however, express about 20 functional performance needs in quantitative terms.[9]

> The F-X will be used throughout the world in varying geographic and climatic environments wherever a threat to the United States exists. Base facilities will range from very austere to extensive base complexes. To provide acceptable operation and reduce requirements for

[9]It specified performance parameters such as takeoff and landing distance, combat radius, maximum speed, ferry range, specific excess power, and instantaneous load factor.

supply support, highly reliable subsystems are required. Military specifications for overall systems reliability must be met. The aircraft design must provide the capability to be maintained in an austere environment within the skill levels available within the Air Force today. AGE requirements should be held to a minimum and be simplified to reduce logistics problems associated with long-range deployments and dispersal of forces to austere operating bases.

We cannot be sure that the F-15 in the field today would be different if a more quantitative expression of operational suitability needs had replaced the citation above because acquisition outcomes are the product of many factors. However, we know that while the F-15 has satisfied most functional performance expectations first expressed quantitatively in its need statement, its operational suitability characteristics do not compare well with qualitative expectations expressed in the paragraph above, although in many respects its characteristics do represent an improvement over those of predecessor aircraft such as the F-4.



SOURCE: ROCs and SONs as noted.

Fig. 5—Quantification of operational suitability characteristics in need-statements has been uneven

The F-15 is one of the most capable fighter aircraft in the world, but it has extensive requirements for maintenance and logistics support. The need for 13 to 18 C-141 loads of flight-line equipment and spares to deploy a squadron to a prepared main operating base detracts from its mobility. And much more equipment and spares are needed to set up at an unprepared base. Failure rates and costs of its avionics require that a minimum of 3 C-141 loads of Avionics Intermediate Shop (AIS) equipment deploy with a squadron to provide a repair capability for avionics line replaceable units (LRUs). This equipment, which requires many highly skilled personnel to operate, is not ideally suited for operations in austere environments, since it requires 4500 square feet of level, air-conditioned floor space.[10]

We found some indications that the relative emphasis accorded functional performance and operational suitability characteristics in need-statements also seems to propagate to other acquisition documents as well. For example, the approved program performance parameters in the *Selected Acquisition Report (SAR)* for the F-15, prepared quarterly by the service program manager for the Office of the Secretary of Defense and the Congress to track program progress, included eight functional performance parameters and only two operational suitability parameters.[11]

The *ROC* for the next Air Force fighter, the F-16, contained considerably more quantitative detail about operational suitability (see Table 2). However, because it was issued after design and flight of the prototype vehicle, and after full-scale development was well under way, it is arguable whether it had much influence on the operational suitability designed into the weapon system, or merely underscored existing design expectations.

The more recent *For Coordination SON for the ATF* is appreciably better than the need-statements that led to the F-15 and F-16. It quantifies many key aspects of operational suitability, and uses more operationally oriented measures—an encouraging start for at least one high-visibility Air Force need.

---

[10]M. B. Berman, *Increasing Future Fighter Weapon System Performance by Integrating Basing, Support, and Air Vehicle Requirements*, The RAND Corporation, N-1985-1-AF, April 1983.

[11]The functional performance parameters were maximum speed, design mission radius, thrust-to-weight ratio, engine thrust, takeoff and landing roll, specific excess power, radar range, and takeoff gross weight. The suitability parameters were mean time between failures and percent operationally ready. The F-15 *SAR's* emphasis on functional performance was typical of most programs. The A-10 *SAR* included eight functional performance parameters and one suitability parameter. See *Selected Acquisition Reports* for the F-15 and A-10.

Table 2

OPERATIONAL SUITABILITY MEASURES
QUANTIFIED IN THE F-16 ROC

Readiness	**Durability**
Flyable rate	Airframe life
Mission reliability	**Mobility**
Mission completion success probability	Support equipment transportable by C-130
Maintenance reliability	**Training**
Contractual mean time between failure	Availability
Maintainability	Mean time to repair
Aircraft turnaround time	Maintenance man-hours per operating hour
Maintenance man-hours per flying hour (total/direct)	Scheduled inspection time
Mean time to return to service	Manning and skill level
Unambiguous fault detection and fault isolation capability by OHMS and BIT	Unambiguous fault isolation capability of ATE
Fault isolation capability to component with technical data and BIT	
3/5 skill level requirement	
Mean time to repair LRU	
Fault isolation of support equipment	
Support equipment calibration level	
Programmed depot maintenance interval	

SOURCE: *F-16 ROC*, 27 February 1976.

Adequacy of Measures

Figure 5's first-order overview of the quantification of suitability needs masks some serious problems in the selection of measures when needs are expressed quantitatively. There is inconsistency in the measures used, in the use of a mix of operational and contractual measures, and in the level of detail—system or subsystem—covered by the measures. For example, system reliability needs are variously stated in terms of break rate (e.g., ATF), or Mean Time Between Failures (e.g., F-16), Removals (e.g., JVX), or Maintenance Actions (e.g., Near Term Manned Bomber and Cruise Missile Carrier Aircraft). At times the terminology implies narrowly defined contractual values (e.g., F-16), while at other times need-statements use broader, more operationally oriented, measures that apply in the field (e.g., Cruise Missile Carrier

Aircraft). Reliability goals are sometimes stated for selected subsystems such as avionics (e.g., JVX), and at other times for the entire weapon system (e.g., FAC-X). Similar problems were identified in the use of measures describing other aspects of operational suitability.

Prioritizing Needs

Today, each major command must prioritize all validated[12] needs (i.e., *SONs*) relative to all its other validated needs and relative to its overall budget submission priorities. However, individual *SONs* have only rarely provided a sense of which of the many needs they state are most important, a particularly critical consideration in the resource-constrained acquisition environment.[13] We found some very limited prioritization of functional performance parameters (e.g., the relative importance of achieving particular specific excess power levels in different parts of the flight envelope) but rarely prioritization across different aspects of functional performance, prioritization of operational suitability characteristics, or the integrated prioritization of suitability and effectiveness needs. The only exception in our review was the *JSOR* for the AMRAAM, which does use an integrated prioritization approach.

Conclusions of Review

Opportunities exist for improving the expression of operational suitability needs in *SONs*. We have argued for the desirability of stating operational suitability deficiencies of existing systems, identifying technological opportunities for enhancing the suitability of weapon systems, and expressing when improved operational suitability is one of the bases for an operational need. There are no intractable obstacles to incorporating these considerations in future *SONs*; the *For Coordination SON for the ATF* is proof of that fact. To ensure that less-prominent *SONs*, not generating the interest of the *ATF SON*, also give prominence to operational suitability issues, may require some changes in mindset and more discipline in the requirements process,

[12]The Requirements Review Group at Air Force Headquarters is the Air Force validation authority for all *SONs* except those below a certain funding threshold.

[13]Explicit guidance calling for prioritization of needs in *SONs* was first issued in 1981 as interim policy direction in advance of publication of a formal revision to AFR 57-1 in 1985, which also directs that needs be prioritized. See *New AFR 57-1 Requirements and Programming Process*, Headquarters USAF, DCS/Research, Development and Acquisition, Requirements, Programs, and Studies Group, 27 August 1981. With very few exceptions, neither *SONs* issued before or after this interim policy guidance have used prioritization to any great degree.

extending from the authors of *SONs* to those ultimately responsible for validating needs. The R&M action memorandum issued in 1984 makes clear the expectations of Air Force leadership in this regard.

This review illustrated the frequent lack of quantification of key operational suitability needs. Ironically, some *SONs* have focused on specifying characteristics of the means for satisfying a need rather than describing the end result desired—e.g., specifying weapon system R&M characteristics while ignoring specification of the end result, which is mission-effective sortie-generation capabilities. Many needs that have been quantified could have profited from the use of better measures. In addition, the priorities attached to achieving particular needs have seldom been expressed. Approaches for identifying needs that require quantification and for developing more descriptive measures are particularly critical to the effective expression of suitability needs.

RECOMMENDATIONS FOR IMPROVING SONs

Identifying Suitability Characteristics to Quantify

Critical factors in deciding which needs to quantify are an understanding of the operating environment and the factors shaping that environment, how an operator expects or hopes to operate in that environment, and how the operator measures system performance. How one might go about identifying which suitability needs to quantify for a potential tactical fighter application is described below.

Future tactical fighter forces will have to be capable of sustaining high mission-effective sortie-generation rates, but *SONs* must say more than that for the R&D community to make sortie-generation rate an important consideration in the design tradeoff process. To help identify the characteristics contributing to the achievement of the sortie-generation force characteristic, one can construct a stylized representation of the sortie-generation process, as shown in Fig. 6. The operator's bottom-line suitability measure is the number of mission-effective sorties that can be flown over time by the force of aircraft, or on a per system basis, the availability of the aircraft for combat, its sortie-generation capability over time, and some measure of the quality of these sorties. These performance-oriented output measures of operational suitability are valued by the user for this mission, measure aspects of readiness, sustainability, and mission accomplishment, and are appropriate for expression in *SONs*.[14]

---

[14]Mission reliability, a partial measure of the quality of the sortie, while an output measure, can also be regarded as a weapon system design characteristic.

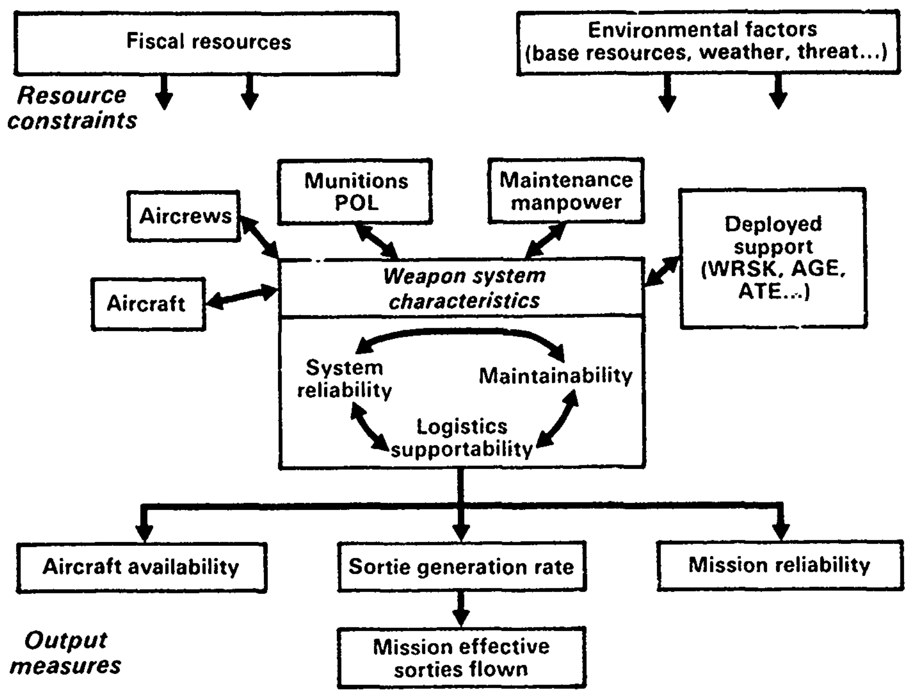


Fig. 6—Factors influencing sortie-generation activity

Constraints on the resources used by the sortie-generation process, shown at the top of the Fig. 6, limit the ability of the operator to generate the desired sorties. The weapon system characteristics shown in the middle of the figure influence the nature and extent of the demands for resources. In principle, different weapon system design concepts possessing different reliability, maintainability, and logistics support characteristics may make different demands for resources while still satisfying the operator's output measure of performance. Design studies must explore these tradeoffs.

In an operational suitability sense, a particular weapon system's attractiveness to an operator depends more on the mission-effective sortie rate it can demonstrate in particular operating environments and the nature and extent of its resource demands than on its inherent R&M and logistics characteristics. The operator thinks in terms of the military mission that must be accomplished with the resources he has at his disposal in a given environment. Consequently, it appears appropriate that *SONs* address applicable constraints such as fiscal resources, manpower, deployed support, and those constraints imposed by the operating environment.

Whether need statements should also express weapon system suitability characteristics (i.e., reliability, maintainability, and logistics supportability) is more problematical. Several factors argue for emphasizing expression of suitability performance outputs and constraints rather than those weapon system characteristics:

1. Such an approach is consistent with current guidance in AFR 57-1 that *SONs* that will result in major acquisition programs should state needs and not solutions.

2. Such an approach yields a more operationally oriented document.

3. Major operating commands preparing *SONs* have a comparative advantage and expertise in expressing operating constraints and operationally oriented measures of desired performance outputs. The development community is in a better position to subsequently determine the R&M and logistics support characteristics required to meet the needed performance outputs.

4. Defining operating constraints and outputs supports the R&D process without unduly constraining the examination of various design tradeoffs of R&M and logistics characteristics.

5. Finally, when *SONs* are drafted, before the start of the Demonstration and Validation Phase (i.e., advanced development), R&M and logistics characteristics have usually not stabilized to the point that one particular combination can with confidence be stated as being most desirable.

Proponents of including more detail about weapon system reliability, maintainability, and logi supportability characteristics argue that it is desirable to give designers an idea of what the ultimate user of the system expects On balance, we believe arguments for specifying the need in terms of the end result desired, using output-oriented measures of military capability, seem stronger than those for specifying the means to that end-result in terms of R&M and logistics support characteristics. In any event, expressions of weapon system R&M and logistics requirements must subsequently be rigorously stated, after design studies have identified the most attractive combinations of those characteristics, which would normally occur sometime during the

Demonstration and Validation phase and certainly no later than the start of full-scale engineering development.[15]

Selecting Measures

Working through the framework of Fig. 6, we considered the appropriateness of various measures for expressing performance-oriented output needs and resource constraints to illustrate the process one must go through to express key suitability needs. These measures include (1) sortie-generation rate, (2) availability, (3) mission reliability, (4) number of aircraft, (5) aircrews, (6) munitions, and petroleum, oil, lubricants (POL), (7) maintenance manpower, (8) deployed support, (9) fiscal resources, and (10) environmental factors.

Sortie-Generation Rate. Perhaps more than any other operational suitability need, a wartime sortie-generation-rate goal expressed in a *SON* should be the subject of tradeoff analyses to determine the implications of different rates on costs, manpower, and other constraints. Such a goal should be regarded as a benchmark for starting design trade studies. Information from the trade studies can then be factored back through the *SON* review process to provide a more informed expression of suitability needs. For completeness, the wartime sortie-generation rate need should be specified in terms of sorties per day over time, mission, and environment, for sortie-generation needs may necessarily be less stringent in more austere operating environments. Expressing an expected average peacetime utilization rate can support life-cycle costing studies.

Availability. Mission-capable rate, the sum of partial and full mission capability, measures availability in terms of the fraction of possessed time that a weapon system is capable of performing any of its designed operational capability (DOC) missions for an assumed utilization rate, mission breakdown (e.g., air-to-air, air-to-ground, etc.), and maintenance operation. The mission-capable rate needed should be defined for each assigned mission. It too should be the subject of tradeoff analyses.

Availability can also be defined in terms of not-mission-capable rate due to maintenance or supply factors, or both, but this seems inadvisable for *SONs*, since only subsequent design studies can identify the most desirable tradeoffs between maintainability and logistics supportability.

Mission Reliability. Mission completion success probability or weapon system reliability is commonly used to express mission reliability. It measures the probability that a scheduled mission will be

---

[15]Section III discusses approaches for expressing and documenting operational requirements for R&M and logistics support characteristics.

completed without experiencing an air abort, ground abort, or other mission deviation that prevents the aircraft from successfully completing all mission objectives. If the need is for a system with a multimission capability, then authors of *SONs* should differentiate among mission reliabilities for each mission as appropriate.

Use of a mission reliability measure has several desirable attributes. It complements sortie-generation goals by partially measuring the quality of the sortie generated. Its expression also forces designers to consider tradeoffs between equipment reliability and redundancy to satisfy mission reliability needs. Such examinations may help identify that level of mission reliability at which acquisition costs begin to increase significantly.

Number of Aircraft. At any given location, the number of aircraft available is obviously an important factor in the number of effective sorties that can be flown. Ideally, however, in a need-statement, it would be preferable to specify the military task that must be accomplished and let the operational suitability and effectiveness of the system determine the force size required to satisfy the need. Although not having an *SON*, C-17 program requirements documentation employed such an approach, specifying a cargo-hauling requirement, letting the number of units needed to satisfy the requirement, at least initially, fall out of design studies. Such an approach may be more difficult when expressing needs that will most likely be filled by a multirole fighter aircraft, however.

The minimum numbers of fighter aircraft needed may be set by the requirement to maintain a simultaneous presence in widely separated locations around the globe, minimum practical deployment and employment unit sizes (trade studies can help in this determination), and force structure considerations. These factors may require specifying the number of units required to satisfy the need in the *SON*, rather than the overall military task that must be accomplished, although the latter approach remains a desirable objective.

Aircrews. Unavailability of aircrews at an airbase can constrain the ability to generate sorties, although generally crew ratios have not been included in *SONs*, and this is not nearly so constraining a factor as are shortages of qualified maintenance manpower. Nonetheless, the utility to the training community of information about the anticipated crew ratio might argue for its inclusion.

Munitions and POL. Although constraints on the availability of these two resources have historically not been noted in *SONs*, increasingly stringent mobility requirements may make inclusion of munitions and POL constraints desirable for future *SONs*. Depending on the operating location and the availability of these kinds of resources, transportation resources required for the continuing resupply of munitions and POL can dwarf the initial transportation requirements to deploy a squadron. Hence, if the operational need is for a system that can operate with a minimum of resources, then limits on the availability of these resources might legitimately be included in *SONs*. Inclusion of such a constraint, expressed in some quantity measure such as tons or gallons per day or equivalent airlift or tanker loads, could influence the selection of a single- or multi-engine design configuration, as well as the kinds of munitions employed and the sophistication of armament subsystems.

Maintenance Manpower. Defining a goal for maintenance manpower levels can stimulate examination of opportunities for reducing manpower requirements through R&M improvements, adjustments in operational concepts, and possible alterations to personnel structures and training approaches.[16] Tradeoff analyses can characterize the comparative effects of each approach and the potential changes in costs and operational effectiveness associated with manpower reductions.

SONs should define limits on the direct maintenance manpower assigned to Aircraft Generation Squadrons responsible for on-equipment maintenance on the flightline, and manpower assigned to Equipment Maintenance and Component Repair Squadrons responsible for off-equipment maintenance in intermediate level (back) shops off the flightline. Table 3 illustrates that today this manpower, spread across 27 Air Force Specialty Codes, accounts for about 81 percent of all maintenance manpower associated with an F-16 wing, and more than half of the wing's total manpower.

In some circumstances, it may be desirable to tailor the manpower limit to focus the development community's attention on particular objectives by selectively including or excluding particular categories of manpower from the limit. For example, to encourage examination of opportunities for streamlining maintenance manpower needs through R&M improvements, one could set goals for limits on manpower in those categories sensitive to changes in R&M, including personnel responsible for the airframe, aerospace ground equipment (AGE), propulsion, and avionics systems.[17] Trade studies could then define the

[16]See Berman (1985) for a discussion of the potential of these three alternatives for reducing maintenance manpower requirements when operating in dispersed modes.

[17]Ongoing RAND research is identifying those categories of personnel most sensitive to changes in R&M characteristics. For example, an across the board doubling of F-16 reliability could reduce an F-16 wing's need for avionics maintenance personnel by 29 percent at three sorties per aircraft per day. Propulsion maintenance personnel needs would drop by 32 percent, whereas loaders, weapons release, and gun servicing personnel needs would decline by only 6 percent, because their activities are not so directly affected by R&M improvements.

Table 3

APPROXIMATE MANPOWER DISTRIBUTION IN A 72 PAA F-16 WING

	Direct Manpower		Overhead Manpower[a]
Maintenance-related wing manpower			
Aircraft generation squadron (contains 3 acft maint units)	666		105
Equipment maintenance squadron	445		47
Component repair squadron	321		44
Deputy chief of maintenance	—		142
	(1432)		(338)
Total maintenance manpower		1770	
Non-maintenance wing manpower			
Wing headquarters	128		
Base operating support	357		
Security	288		
Real property maintenance	33		
Medical	54		
Operational squadrons	100		
Total non-maintenance manpower		960	
Total, all wing manpower		2730	

SOURCES: RAND research in progress, *USAF Cost and Planning Factors*, AFR-173-13.
NOTES: 3 1.7-hour sorties per day; 7-day surge; 2 12-hour maintenance shifts per day; Combat Oriented Maintenance Organization (COMO).
[a]Administration, supervision, command.

potential costs of obtaining the R&M improvements and the potential benefits and liabilities of the smaller maintenance workforce.

The number of personnel involved in munitions assembly, missile maintenance, gun servicing, the loading of munitions, and weapons release functions are more tenuously influenced by changes in R&M characteristics, although the amount of manpower performing these functions can be influenced by the operational effectiveness of systems proposed to satisfy the need.[18] Manpower limits can be expressed on a per aircraft basis or totals for a wing. Accompanying descriptions of the wartime and peacetime operational settings are particularly

[18]Ongoing RAND research is examining the influence of R&M improvements on maintenance manpower requirements.

important when expressing limits on manpower, because manpower levels are strongly influenced by sortie rates, shift policies, and maintenance concepts.

Deployed Support. The need for mobility in future systems suggests the desirability of keeping deployed support requirements to a minimum. *SONs* can, but seldom have, set aggregate limits on deployed support in the context of expressing mobility needs. (Only the ATF *SON* has quantified such a limit.) They have at times qualitatively discussed packaging, handling, and transportability needs without quantitatively addressing particularly critical aspects of mobility.

Mobility attributes include reaction time, ease of movement, transportation resources required for movement, rapidity of movement and establishment of a combat capability, and flexibility in choice of operating locations. Mobility depends both on functional performance characteristics and the nature and extent of support resources needed for deployment.

Factors such as ferry range, the presence or absence of an inflight refueling capability, flight aids, and cruise altitude contribute to ease of movement. Cruise speed influences rapidity of movement. Flexibility in the choice of operating locations is influenced by required runway lengths, widths, and surface requirements, as well as combat radius.

The nature and extent of support that needs to be deployed influences the time required to prepare and load support elements after notification to deploy, the transportation resource requirements for a deployment, and the time to set up and establish a combat capability at the destination. Expression of each of these factors could enhance the descriptiveness of expressions of mobility needs in *SONs*. Inclusion of limits on time factors associated with deployment can encourage design studies of support elements that do not require extensive and time-consuming disassembly for loading and subsequent reassembly after unloading. No *SONs* we reviewed included such time factors.

Setting limits on transportation resource needs for deployment constrains the amount of deployed support and, in so doing, can encourage studies evaluating the attractiveness of improved R&M, built-in support features, alternative operating concepts, and modified personnel structures in enhancing mobility. Such studies might show, for example, how mobility, aircraft gross weight, and consequent functional performance and cost are influenced by various shifts to built-in support on the fighter aircraft.

There is no single right answer to the question of what aspects of deployed support should be subject to a transportation constraint in a

SON. Depending on the environment to which a unit deploys, all or some of the things shown in Table 4 will need to be deployed for support. However, as we discuss below, not everything shown in Table 4 appears appropriate for inclusion under a deployed support constraint in a weapon system need-statement.

Virtually every need-statement for a tactical aircraft that has a mobility mission should include those things listed in category 1 of Table 4. When deploying to an established base in a theater such as

Table 4

MAJOR CATEGORIES OF DEPLOYED SUPPORT

Deployed Support Category	Description/Comments
Category 1	
Personnel	Maintenance, officers including pilots, supply/clerical, security specialists, medical corpsman, quality control, maintenance control
War Reserve Spares Kit (WRSK)	For specified time period and sortie rate
Ground Support Equipment (GSE)	Flight line and intermediate GSE
Other	Spare engines, tools, equipment, multiple ejection racks, triple ejection racks, munitions build-up items, munitions racks, office supplies, small arms and ammo, communications equipment, . . .
Category 2	
Petroleum, oil, lubricants (POL)	
Munitions	Built-up or not built-up
Category 3	
Housing/facilities, utilities	Harvest Bare, Harvest Eagle, . . .
Food	Field rations or more elaborate meals
Water	
Antiaircraft protection	Stinger surface-to-air missiles
Medical supplies	
Command, control, communications centers	Tactical Air Control Centers, Control and Reporting Centers, Control and Reporting Posts, Forward Air Control Posts, . . .

SOURCE: Berman (1985).

NATO, those things essentially encompass the deployed support need. For an F-16 squadron having an air-to-ground mission, these four classes of support account for the equivalent of approximately 20 C-141B transport aircraft loads (see Table 5), with ground support equipment accounting for half of the total airlift needed. The development community can influence the amount of this deployed support through weapon system design initiatives. Moreover, if *SONs* include manpower constraints, then those constraints will also influence the amount of other kinds of deployed support, such as housing, food, and water, which may be important in more austere operating environments. Aside from this indirect but important influence on deployed support falling in category 3, other efforts to reduce the amount of deployed support needed from category 3 seem best addressed in initiatives separate and distinct from tactical aircraft need-statements, because that category of support would not fall under the purview of the aircraft weapon system development program designed to satisfy the need.

The appropriateness of including or excluding POL and munitions from an airlift (or other transportation) constraint is more problematic and is seemingly closely tied to the intended operating and employment scenario associated with the operational need. In deploying to an established base in NATO with direct access to these two resources, inclusion of POL and munitions seems inappropriate. Even in a tactical dispersal in the NATO theater, distances to these resources would not be extreme and resupply would normally be accomplished through ground transport.[19] Hence, it may not be desirable to include POL and munitions as part of a transportation constraint in this situation either.

In more austere environments, resupplying POL and munitions can be a considerable burden. If the user deems constraints on the consumption rates of these resources of compelling importance in satisfying the need for operations in such environments, such that he desires that they influence designed effectiveness and fuel efficiency directly, then it may be desirable to include constraints on these resources directly, as was noted earlier. If constraints on *resupply* levels for POL and munitions are specified, they should be stated separately from constraints on *initial deployed support* falling in the first category of Table 4, because of the different character of initial and recurring support.

Fiscal Resources. Ultimately, fiscal resource constraints influence the availability of all the other resources shown at the top of Fig. 6 that contribute to the sortie generation process. However, it would be

[19]See Berman (1985).

Table 5

AIRLIFT REQUIREMENTS FOR INITIAL DEPLOYMENT OF AN F-16 SQUADRON WITH A PRIMARY AIR-TO-GROUND MISSION

Category	Weight (lb)	Equivalent C-141B Loads	Percent of Airlift
Ground support equipment	446,000	10.1	51
Other	254,000	5.8	29
War Reserves Spares Kit	37,000	0.8	4
	737,000	16.7	84
Personnel (552 people)	144,000	3.2	16
Total	881,000	19.9	100

SOURCE: Developed from F-16 Unit Type Codes (UTC). See Berman, R-3304/1-AF.

NOTES: 24 aircraft in the squadron; 30-day operation; maximum wartime sortie rate; deployed support includes munitions build-up package, associated personnel, and GSE; Table 4 explains "other" equipment; 22 tons or 175 persons per C-141B load.

almost impossible to state some aggregate fiscal resource constraint on the sortie-generation process as a whole in a *SON*, and in any event would probably do little to shape R&D activities. In contrast, addressing affordability issues could materially influence R&D activities.

There is appreciable institutional reluctance to address affordability considerations in *SONs*, even though costs are ultimately a significant determinant of the quality and quantity of systems acquired to satisfy needs. None of the aircraft *SONs* we reviewed addressed affordability issues in a satisfying quantitative sense. One justification advanced is that cost constraints should not cloud a description of the operational characteristics really needed to counter a threat. This view holds that any compromises in those characteristics required because of fiscal and technical realities should be addressed only outside the context of the *SON* as the requirements process proceeds, while the *SON* remains an accurate reference for the expression of any threat-driven needs. A second, more pragmatic, unstated reason for the reluctance may be an institutional desire to avoid offering up a quantitative cost goal as a focal point for debate that could adversely affect the validation of a need.

Although these arguments may have some merit, affordability considerations seem too important in shaping R&D activities to exclude from *SONs*. Ideally, *SONs* should include cost constraints for initial and recurring costs, such as a constraint on unit flyaway cost and recurring operating and support cost for needs expected to be satisfied by an aircraft. Expressing a credible range of costs rather than a point estimate can convey the affordability constraint without stifling the full exploration of design alternatives. Expressing the numbers of aircraft (or other systems) needed (as discussed earlier) can explain the quantity dimension of cost.

Environmental Factors. *SONs* should explain the operational environment or setting within which needs must be satisfied. Descriptions must be tailored to the particular need using a mix of quantitative and qualitative statements. Figure 4 noted major subject areas requiring treatment. The forthcoming Air Force regulation governing the preparation of the *System Operational Concept* document, AFR 55-24, *Concepts, System Operational Concepts*, and AFR 66-14, *Equipment Maintenance, Equipment Maintenance Policies, Objectives, and Responsibilities*, can provide useful information for compiling an adequate description of operational settings.

Descriptions of maintenance and basing concepts will be particularly important for needs that will be satisfied by tactical aircraft that will at times operate away from main operating bases in less than wing-size units. Some of the more important elements of the basing environment include the characteristics of the landing and taxiway surfaces, the kinds of bases (main operating bases, other operating bases, unimproved operating sites such as bare bases or forward operating locations), the size of aircraft units operating from those bases, and facilities already at the base, such as stands, jacks, powered support equipment, fuel, munitions, utilities, shelters, etc., that are exclusive of the deployed support package.

Detailed and standardized descriptions of various kinds of Air Force operating locations could serve many recurring activities in support of the acquisition process, including preparation of *SONs*, *SOCs*, Requests for Proposals, Statements of Work, and contracts. Formal publication of descriptions of these reference or generic bases could provide a common basis for discussing various support issues associated with the satisfaction of future needs.[20]

Maintenance concepts will usually vary according to the type of operating location. Some of the more important things *SONs* should describe include where (e.g., flight line, intermediate shop, depot) and how (e.g., remove and replace, remove, repair, and replace)

[20]The forthcoming Air Force Pamphlet AFP 57-9, *Defining Logistics Requirements in SONs*, could be an appropriate place for such descriptions.

maintenance will be accomplished, unit self-sufficiency times (duration of operating capability without external support), deployed unit sizes, and the organization of maintenance units for each kind of location.

Using effective measures to express key suitability needs depicted in Fig. 6 is an important step in improving need-statements. Prioritization, discussed below, can then give a sense of what needs are most important.

Prioritizing Needs

Although developing a documented list of prioritized suitability needs presents methodological and institutional difficulties, doing so has the potential for beneficially supporting activities from the initial expression of a need to the final stages of testing and the initial fielding of the system satisfying that need. Setting priorities could help the R&D community structure design-tradeoff analyses to yield the kinds of information the Air Force needs to identify desirable design solutions. Prioritized needs could subsequently be used as an evaluation aid during source selection. They could also potentially be used to support decisions during resource-constrained development and test activities, and to support the prioritization of deficiency corrections during and after testing.[21]

Ideally, needs selected for prioritization should be broad in scope, preferably quantitative, but limited in number to avoid diluting the impact of the prioritization. One would expect both the list of needs and their priorities to differ across mission applications. Candidates for prioritization for a tactical aircraft need might include such aspects of operational suitability as weapon system availability, wartime sortie-generation capability, mission reliability, airlift resources needed to deploy a squadron, and number of direct maintenance personnel, supplemented by unit flyaway cost and recurring operating and support cost.[22] A need statement for a strategic aircraft mission capability would probably share some of the same measures of operational suitability, but they quite possibly would occupy a different position in the prioritization structure. Availability might be of higher priority for a

[21]Recent changes in the Air Force deficiency correction process direct that all service reports (which identify deficiencies) be prioritized in programs involving operational test and evaluation. A variety of techniques have been used to prioritize deficiencies. Prioritization of needs in a *SON* could help in these processes of ranking deficiencies. See *Interim Guidance to T.O. 00-35D-54, Section V*, 16 April 1985, issued by the Maintenance Policy Division, Director of Maintenance and Supply, Deputy Chief of Staff for Logistics and Engineering, Headquarters USAF.

[22]Although unstated here, flight safety, and personnel safety in general, understandably always occupy a preeminent position in aircraft design priorities.

strategic aircraft mission need, whereas sortie generation rate might be comparatively more important for a tactical aircraft mission need.

Establishing priorities would inevitably involve use of a mix of analytic techniques and subjective professional judgments.[23] As experience grows in dealing with methodological and institutional difficulties in documenting priorities, the Air Force could set a long-term objective of insisting on inclusion of an integrated prioritization of operational effectiveness and suitability needs in each *SON*. Our research identified o. ' one program that has done so. the AMRAAM air-to-air missile development. The priorities established in the *Joint Service Operational Requirement* for that system guided contractor efforts during the AMRAAM design competition and were used during source selection evaluation.

Improving Policy Guidance

Having adequate policy guidance to help authors of *SONs* select appropriate measures to express operational suitability needs is particularly important. Air Force personnel responsible for drafting requirements rotate through operating command requirements organizations frequently, and corporate knowledge can be lost in the process, although civilian personnel are an important source of continuity. We found that, although requirements documents must always be reviewed by higher authorities, it is not unusual for the task of drafting such documents to fall on a few junior personnel. Good policy guidance will never replace experience, but seemingly it can help authors of need statements, particularly when they are operating under time pressure, as they often are. With many needs competing for inclusion in *SONs*, guidance about what aspects of operational suitability should be addressed in need statements (see below) might also promote the development of more balanced documents that address both effectiveness and suitability needs.

Guidance for expressing operational needs generally falls into one of two categories, that guidance dealing with administrative or procedural aspects of documenting needs, and guidance dealing with the substantive content of *SONs*. The Air Force has undertaken a number of initiatives during the past several years to improve the administrative processing of *SONs*, culminating with issuance of a new version of AFR 57-1, *Operational Requirements, Operational Needs*, in the summer of 1985, hence we focused on guidance about the substantive content of

[23]There is a growing literature and interest in methods for quantifying subjective judgments. See, for example, Crawford and Williams (1985).

SONs, and in particular, guidance for selecting operational suitability measures.

Based on our experience in exploring approaches for expressing suitability needs for tactical fighter aircraft, we believe such guidance should identify and define measures, and explain the rationale for selecting particular measures to define suitability needs for different types of system applications. As Table 6 indicates,[24] guidance is available that identifies and defines many measures, but comparatively limited guidance is available about possible strategies for helping authors of need statements decide which measures to select to assemble a cohesive set that characterizes the major suitability needs for a particular mission element need. Guidance for doing so would have to link the use of particular measures more explicitly to particular military missions and to the way in which an anticipated system would be used. The conceptual approach presented earlier in this section for identifying needs to quantify, and measures to use for expressing tactical fighter needs, illustrates an approach that moves beyond lists of measures and their definitions. Similar guidance for an air-to-air missile might explain how various measures characterize important aspects of such a missile's life cycle, including dormant storage, captive carry, and employment, as illustrated in the AMRAAM life-cycle profile shown in Fig. 7.

The most recent revision of AFR 57-1 focuses more on the administrative aspects of processing needs than their substantive content. The most likely vehicle for including material that discusses strategies for selecting measures would be a new pamphlet, AFP 57-9, being circulated in draft form at this writing, that is designed to assist authors in developing the substantive aspects of suitability needs expressions.

The effective expression of needs is a necessary first step in the process of developing more operationally suitable systems, but effective follow-up actions with respect to requirements are needed to support subsequent activities of the acquisition process. The next section describes some ideas for improving the substance and documentation of these operational requirements.

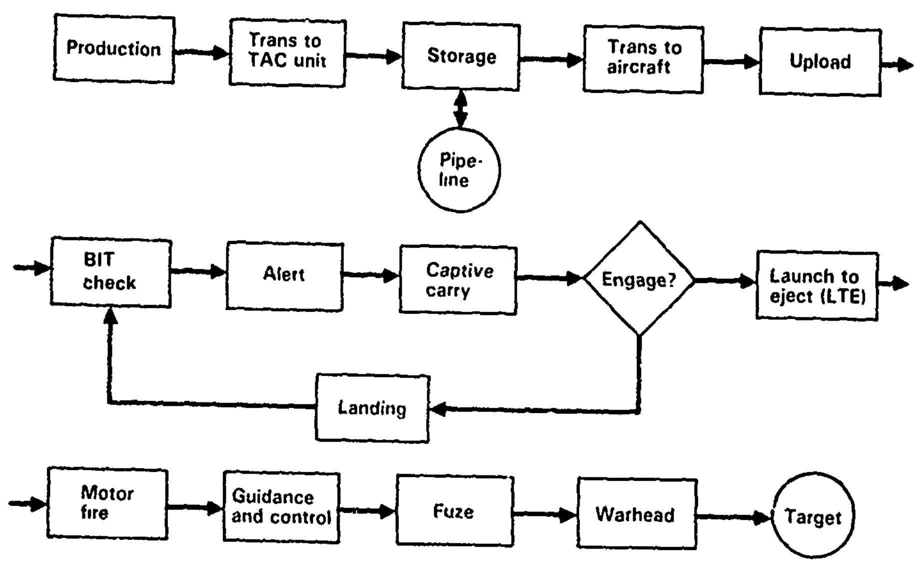
---

[24]Table 6 lists Air Force-wide and DoD-wide documents. Other command and separate operating agency documents contain useful material on measures for expressing suitability-related needs. See, for example, *Logistics Assessments*, Air Force Operational Test and Evaluation Center, AFOTEC Pamphlet 400-1, 29 February 1984.

Table 6

GUIDANCE FOR SELECTING OPERATIONAL SUITABILITY MEASURES

Source	Title	Identifies Measures	Defines Measures	Discussion of Selection Rationale	Comments
DoDD 5000.40	R&M	Very limited	No	Very limited	Basic policy guidance from DoD
MIL-STD-721C	Military Standard, Definitions of Terms for R&M	Yes	Yes	Very limited	Supports 5000.40
AFR 800-18	Acquisition Management, AF R&M Program	Yes	Yes	None	AF implementation of 5000.40 AF approved standard measures No diagnostic measures
AFP 57-9	Defining Logistics Requirements in SONs	Yes	Yes	Some	Incorporates guidance from many AF regs Includes some measures not in AFR 800-18
AFR 57-1	Operational Requirements, Operational Needs	Very limited	No	Very limited	Emphasizes administrative aspects Refers to AFR 800-18 for measures guidance



SOURCE: AMRAAM Program Office briefing, 1983.

Fig. 7—AMRAAM life cycle profile

III. EXPRESSING OPERATIONAL REQUIREMENTS

As user's needs are refined, extended, and expanded as part of the interactive process between the user and the development community, they evolve into operational requirements. These expressions of requirements provide more details about the desired characteristics of the proposed solution to the user's need, and are a critical ingredient in the system acquisition process. Program offices use operational requirements to develop the contractual requirements specifications that govern the development of the system designed to satisfy the operational need. The operational test community uses operational requirements to develop test plans and to evaluate whether performance demonstrated during testing meets the effectiveness and suitability requirements of the user. Decisionmakers use these results to gauge whether programs are making satisfactory progress. Supporting commands use operational requirements to plan for the eventual introduction of the system into the inventory. When the operational requirements process breaks down, it can have wide-ranging ramifications.

Through review of operational requirements documentation and discussions with program offices and operational testers, we found indications of some shortcomings in both the expression of operational suitability requirements and in their documentation. We briefly characterize some of the shortcomings in expressions and their effects, and identify some of the more prominent contributing factors. We then suggest some approaches for selecting measures that describe the interactions between R&M and logistics support characteristics influencing the supportability of weapon systems. We then identify some prescriptive measures for current requirements documentation procedures.

SHORTCOMINGS OF OPERATIONAL REQUIREMENTS

In visiting a cross-section of six tactical missile and aircraft program offices involved in activities ranging from advanced development to production, we learned how operational requirements that are poorly stated and documented can complicate the development process, and ultimately influence the ability of a developed system to satisfy an operational need.[1] Problems involved traceability, the meaning and appropriateness of requirements measures, the introduction of new operational requirements after contracts had been let, omissions, and the overprescription of requirements.

In one instance, documentation of a requirement that became a major issue at an AFSARC review was so vague that the operating command responsible for writing it in the first place could not reconstruct its origin, meaning, and rationale.

Other requirements lacked operational meaning. Operational reliability requirements for one fighter aircraft were very narrowly defined, encompassing only inherent design failures, failing to quantify the maintenance burden the Air Force experiences due to failures induced by maintenance, and particularly those failures classified as "no defect" events.

Inadequate understanding of planned operating concepts can also contribute to the development of inappropriate requirements. Requirements for one air-to-ground precision guided munition failed to distinguish between the unique requirements for training rounds and for tactical (live warhead) rounds, the former frequently being carried in captive flight on the aircraft, the latter usually remaining in dormant storage except for occasional "g[illegible]erations" (i.e., removals from storage for loading and testing but not flight). Recovery from this oversight contributed to a delay in the start of high-rate production.

The use of mean time between maintenance actions as the same reliability measure for both training and tactical missiles is also questionable, since the inspection interval set by policy for the tactical missile strongly influences when failures will be noted for it, whereas the more constant use of the training missile makes MTBM a more appropriate measure of its reliability.

In some cases, operational requirements for important aspects of system performance were not stated until after contracts were let, forcing program offices to try to meet the new requirements without significantly restructuring the contract already signed. This occurred in a program that began before the concept of baselining had been fully implemented to deal with such situations.[2]

---

[1]Visits to the six program offices were adequate to obtain insights and indications of problems, but not conclusive evidence that those problems occur frequently in other programs.

[2]When signed by key participating organizations, The *Program Baseline Contract* documents the scope of an acquisition program in such areas as system configuration and performance, operating concept, support/maintenance approach, T&E program, training approach, facilities, schedule, and program costs. It is designed as a management tool to control changes to acquisition programs. See *Acquisition Program Baselining*, AFR 800-25, 30 November 1984.

SOURCES OF SUITABILITY REQUIREMENTS PROBLEMS

Technical, operational, and institutional factors can contribute to shortcomings in operational requirements. The breadth of contributing factors means that no single Air Force organization can address the entire issue. The factors encompass some subject areas beyond the scope of this particular study, but we identify and briefly describe each.

Technical Factors

Equipment Complexity. To meet the threat, aircraft subsystems such as combat avionics are designed with many functions, with each function having a variety of operating states. Trying to characterize the reliability being sought from such systems can be very difficult, because they often suffer performance degradation in one or more of their subsystems instead of suffering the catastrophic failures that conform better to classical reliability measures.

Methodological Limitations. Traditional methodologies accept current basing and support systems as "givens" when assessing the effectiveness of alternative future air vehicle designs and do not fully integrate operational suitability considerations into the design tradeoff process. Development and use of methodologies that more explicitly consider effectiveness and suitability tradeoffs could result in more balanced statements of requirements.[3]

Operational Factors

Mission Complexity. Most tactical aircraft are designed to fulfill a variety of missions. The level of performance needed from particular items of equipment to accomplish a mission successfully can vary from one mission to another. For example, a particular level of radar detection degradation on one kind of mission—rendezvousing with a large target such as a tanker—may not be considered critical, but on another mission—detecting a low-flying, small cruise missile—the same degradation may result in a mission failure. Hence, the mission mix, and relative frequency with which those missions are practiced, can shape perceptions about the reliability of the aircraft. Developing a set of measures that provide an adequate picture of an airplane's reliability in the face of such mission complexity is difficult.

Inadequate Field Data. The aforementioned equipment and mission complexity complicates the ability to collect data routinely for understanding day-to-day use and experience with existing systems, the

[3]See Berman (1985) for a description of such a methodological approach.

about operational needs, but also makes it harder to develop and sustain the institutional memory and analysis expertise needed for generating and expressing new requirements. This increases the importance of effective communication between the operating commands and the development community, the operational test and evaluation community, and other organizations having analytical expertise that can contribute to the development of requirements.

The aforementioned factors contributing to shortcomings in operational suitability requirements statements encompass some issues beyond the scope of this study effort, although as noted above, other RAND research is addressing some of the factors, while other factors are being addressed through Air Force initiatives. We focus on the selection of requirements measures and approaches for improving the documentation of requirements.

SELECTING BALANCED MEASURES

Amending existing policy guidance to illustrate and organize the relationships among key suitability measures could help authors of operational requirements statements select balanced sets of measures for operational suitability. By balanced, we mean measures that (1) address generically important aspects of operational suitability, and that (2) recognize and measure the interactions between R&M and logistics support characteristics that influence the supportability of weapon systems. Expressions of operational requirements for weapon system R&M and logistics support characteristics should be formulated when the most attractive combinations of these characteristics become evident through Demonstration and Validation Phase design studies. These requirements can then complement the expression of resource constraints and output-oriented measures of military capability in the *SON*.

Figure 8 codes each measure according to the aspect of operational suitability it measures and locates it in the R&M and logistics support tradeoff space according to whether it is influenced by one of those characteristics, or some combination of two or three of them.[5]

[5]The meaning of the acronyms used in Fig. 8 can be found in the Glossary. Definitions of the measures and acronyms can be found in *Acquisition Management, Air Force Reliability and Maintainability Program*, AFR 800-18, 15 June 1982; *Defining Logistics Requirements in SONS*, AF Pamphlet 57-9, forthcoming; *Equipment Maintenance, Equipment Maintenance Policies, Objectives, and Responsibilities*, AFR 66-14, 15 November 1978; *C-17 Preliminary Design Review (PDR), Reliability and Maintainability*, Douglas Aircraft Company, McDonnell Douglas Corporation, 7-17 May 1985; *System Specification for C-17 Airlift System*, Douglas Aircraft Company, McDonnell Douglas Corporation, Specification MDC S001, Code Ident 88277, 1 April 1983.

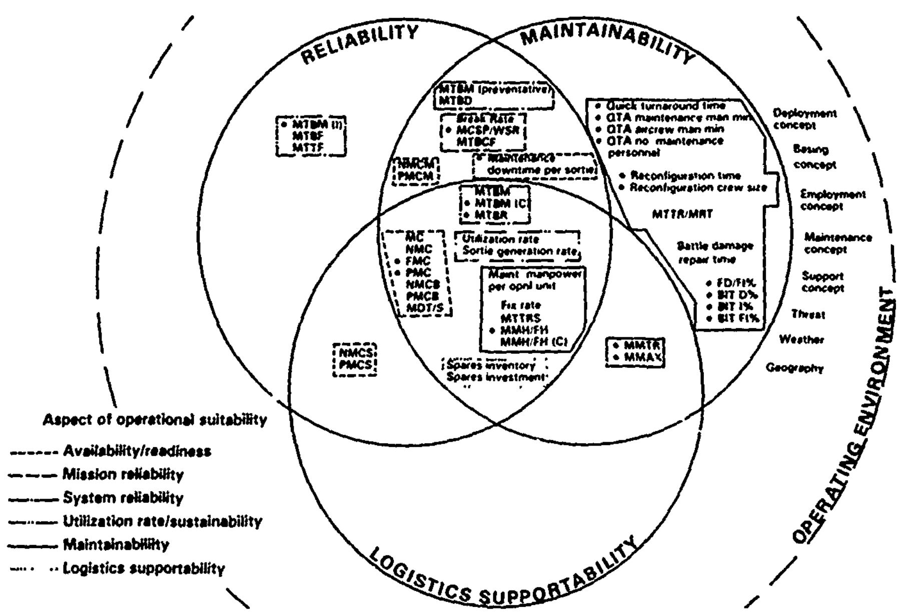


Fig. 8—Relationship of R&M and logistics support measures

Organizing measures in this fashion could complement existing guidance that usually just lists and defines the measures without noting their interrelationships.[6]

The construction of Fig. 8 illustrates that the operating environment influences virtually all aspects of operational suitability. Description of the operating environment is therefore critical for providing the setting within which the operational requirements must be achieved.[7]

It is important to specify measures for inherent characteristics (depicted in the non-intersecting portions of the diagram) (e.g.,

[6]Although basically limited to the development of availability requirements, guidance published by the Navy shows analytically the relationships among some key R&M and logistics support characteristics that influence availability. See *Operational Availability of Weapon Systems and Equipments: Definitions and Policy*, Chief of Naval Material, NAVMAT Instruction 3000.2, 21 January 1981. Draft Air Force Pamphlet 57-9, *Operational Requirements, Defining Logistics Requirements in Statements of Operational Need*, although covering a broader range of suitability measures, presents a less detailed view of analytic relationships than the Navy instruction.

[7]See *Concepts, System Operational Concepts*, AFR 55-24, forthcoming, for guidance about describing operating environments.

characteristics that are a pure reflection of a system's reliability), but specifying measures that capture the interaction and dependencies among characteristics (depicted in the intersecting portions of the diagram) is even more important, because it can encourage a more balanced view of weapon system design characteristics.

A system can have good inherent reliability but create an appreciable maintenance burden for the Air Force if its maintainability is poor because of numerous erroneous failure indications prompted, for example, by a poor fault isolation system. For that reason, a reliability measure such as Mean Time Between Maintenance (inherent failures only) [MTBM(I)] as used in the F-16 program and shown on the left of Fig. 8, has only limited operational meaning. In contrast, Mean Time Between Maintenance (all corrective events) [MTBM(C)], shown in the center of Fig. 8, captures not only inherent reliability, but also the performance of the fault isolation (FI) system (e.g., no defect removals brought about by poor FI), and whether a system has been spared properly (e.g., extra removals due to cannibalization).[8]

Figure 8 can stimulate thinking about interactions and tradeoffs among operational suitability parameters, but to fully explore those relationships requires more sophisticated tools such as current (e.g., Dyna-METRIC, LCOM, etc.) or new sortie operations models that can analyze aircraft needs for support on airbases.[9]

Some progress is evident in the use of operational requirements measures that reflect the interactions of weapon system characteristics. The C-17 program is using MTBM (corrective) as an operationally oriented contractual measure of reliability that cuts across the R&M and logistics supportability parameter space (in the diagram the C-17 measures have dots to their left). The increasingly integrated nature of systems on modern fighter aircraft, and the design challenges associated with developing adequate fault isolation systems for such aircraft, make it especially important that requirements measures capture the maintenance burden due to not only inherent, but also induced and no-defect, failure events.

[8]The value of any R&M measure depends on accurate objective reporting of failures and maintenance actions. Other RAND research suggests that operational judgments about when to report apparent failures and when to request maintenance at times compromise the value of R&M measures such as MTBM.

[9]See, for example, Pyles (1984).

DOCUMENTING OPERATIONAL REQUIREMENTS

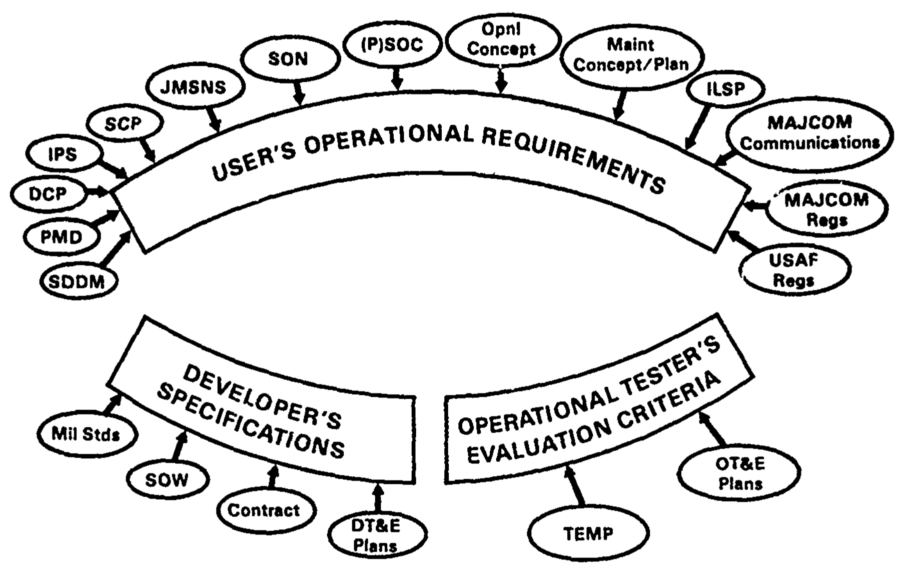
Documentation Problems

Improved procedures for formally documenting operational suitability requirements could complement improvements in the substantive expression of requirements. The fragmentation of operational requirements across many sources and the difficulty in correlating key operational, contractual, and test requirements are basic documentation problems.

Insistence on formal documentation of operational requirements can encourage more discipline in the requirements process. Formal requirements documentation gives program offices a basis and justification for developing contractual requirements and for obligating financial resources. It is the operational test community's basis and justification for planning and conducting tests and evaluating how well a system meets the user's needs. Developers and testers can be put in an untenable position if they must proceed in the absence of documented requirements—particularly if, in the perception of the user, those undocumented requirements change as a system moves through development. The foregoing is not an argument for never changing requirements once set, but rather for documenting requirements both initially and when they must be changed to respond to a changing threat or other factors.

In contrast to the *Statement of Operational Need* document—the vehicle by which operational needs are conveyed—no single document, in practice, serves as a unified unambiguous source of Air Force operational requirements. Various documents serve as primary original sources of requirements, direct changes in requirements, restatements of requirements, or translations of them into contractual specifications or test evaluation criteria (see Fig. 9). (See Glossary for definitions of acronyms in Fig. 9.)

Initially, requirements may be stated in the *SON* by the user, and some occasionally appear in the *Justification for Major System New Start (JMSNS)* document, prepared by the Air Staff, which is needed to initiate funding of a major new system. The user refines and expands needs and requirements in the *Preliminary System Operational Concept (PSOC)* document, and the *SOC* that follows. *Operational Concept* documents prepared by the user occasionally express requirements in broad terms. The user develops the *Maintenance Concept* document for incorporation in the *PSOC* and *SOC*, and subsequently refines it, in conjunction with the weapon system program office, to develop a *Maintenance Plan*, which is incorporated in the *Integrated Logistic Support Plan (ILSP)*.



SOURCE: Documents as noted.

Fig. 9—Documentation of operational requirements is fragmented

Formal or informal communications from the operating command that will ultimately use the system, and operating command or headquarters regulations can also serve as original sources of requirements. The Secretary of Defense can direct changes in operational requirements using the *Secretary of Defense Decision Memoranda (SDDM).* Similarly, the Air Staff can use *Program Management Directives (PMD).* The Defense Systems Acquisition Review Council (DSARC) can direct changes using the *Decision Coordinating Paper (DCP).*

The other documents shown in the lower half of Fig. 9 restate or translate operational requirements in varying degrees to satisfy a variety of acquisition objectives. The development contract translates operational requirements into contractual specifications. The various test plans—both the *Test and Evaluation Master Plan (TEMP)* and detailed development and operational test plans—translate requirements into evaluation criteria to be used during testing.

The multiple sources that document operational requirements are issued by different organizations, at different times, to satisfy different objectives, and the timing and degree of updating varies. This creates the potential for introducing inconsistencies and uncertainties about operational requirements over time and makes it extremely difficult to trace a requirement from its origin to its final manifestation as a test requirement. The results shown in Fig. 10, developed by a joint Air Force Operational Test and Evaluation Center (AFOTEC) and Air Force Systems Command (AFSC) working group, illustrate numerous inconsistencies in requirements documentation for one missile program.[10] The examination of this program's requirements documentation was applied retrospectively after the requirements inconsistencies had contributed to delaying an Air Force System Acquisition Review Council (AFSARC) decision on the program's readiness for production. Reviews of documentation in other programs identified similar inconsistencies.

	DOCUMENTS						
Title of Recommended Change	Required Operational Capability	System Operational Concept	Executive Program Summary	Program Management Directive	Specification	Baseline	Test and Evaluation Master Plan
Mission Reliability		X	X	X	X	X	X
Availability		X			X	X	X
System Reliability		X	X		X	X	X
Weapon System Reliability		X	X			X	X
QRA Launch Times	X	X	X		X		X
Enroute Launch Times	X		X		X		X
Emplacement Time		X			X	X	X

X - Changes made to bring all documents into agreement.

SOURCE: AF Operational Test & Evaluation Center.

Fig. 10—Requirements documentation is inconsistent

Improving Documentation Procedures

Actions to deal with the inconsistencies across requirements documents are needed. The aforementioned working group developed a management tool called the *Baseline Correlation Matrix (BCM)* to iden-

[10]This working group, formed by the commanders of the two organizations in 1983, was charged with developing improved approaches for achieving common baselines between users, developers, and operational testers about operational requirements, contractual specifications, operational test criteria, and evaluation methodologies to ensure more focused and productive milestone reviews.

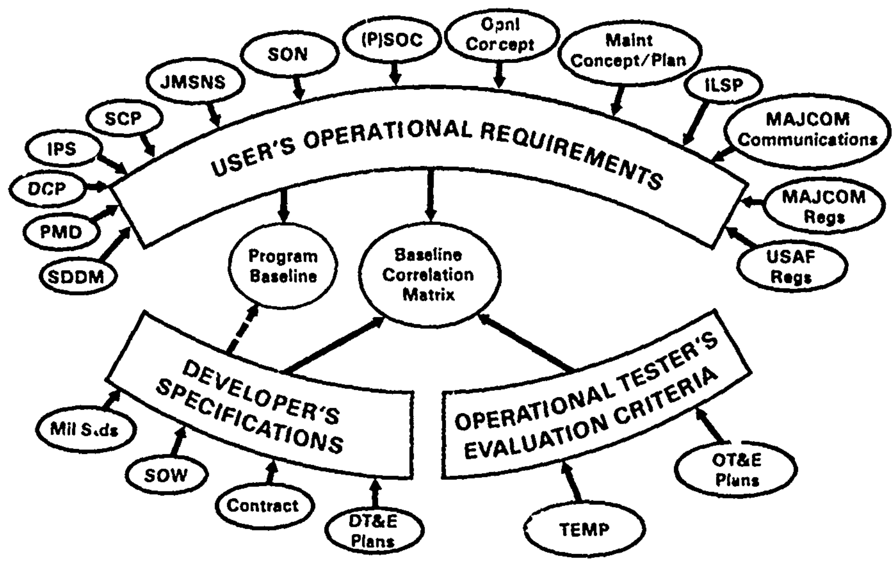
tify inconsistencies needing correction. The BCM, developed jointly by the developer, user, operational tester, and supporting organizations, is not a primary source document for requirements, but rather correlates in one document (1) operational requirements drawn from *SONs*, *SOCs*, *PMDs*, etc., (2) specifications from development contracts, and, (3) evaluation criteria drawn from *TEMPs*, detailed test plans, etc. When inconsistencies are identified in requirements documentation in developing the *BCM*, the organization responsible for the particular requirements document must take the necessary actions to incorporate the needed changes. Ideally, compilation of the *BCM* should begin during the concept development phase and it should be updated through full-scale development as necessary.

At this writing, the *BCM* is being applied on a case-by-case basis—currently to 18 programs—but its use has not yet been institutionalized across the Air Force. It would seem desirable to do so because it treats an important part of the requirements documentation problem by correlating the three critical ingredients of requirements information, as depicted in Fig. 11. Its early development could also contribute to the *Program Baseline*, which incorporates operational requirements, contractual specifications (for information purposes only), and cost, schedule, and related acquisition program information. At this writing, the Air Force had not decided whether the *BCM* should be formally integrated with the *Program Baseline*.

The *BCM* is not an original source document for operational requirements, but reflects and references operational requirements stated in other documents. Over the long term, it may be desirable to consider developing a more unified source of operational requirements, just as the development and production contract serves as an unambiguous source of contractual requirements for the program office and the contractor. This document would consolidate operational requirements in a single document, including those specified by the user, and those directed by other authorities.

The functional requirement for such a document does not necessarily imply a need to create a new acquisition document to add to the many that already exist. The *SOC* (and the document that precedes it, the *PSOC*) in principle, if not in practice, has many of the attributes needed to fulfill the role. It is an original source document for operational requirements developed by the user. Its format is structured to address most major aspects of operational suitability, and it is an active document that is supposed to be progressively updated as a system moves through the acquisition process.[11] How it relates to other

[11]See AFR 55-24 for the structure and revision policy of the *SOC*.



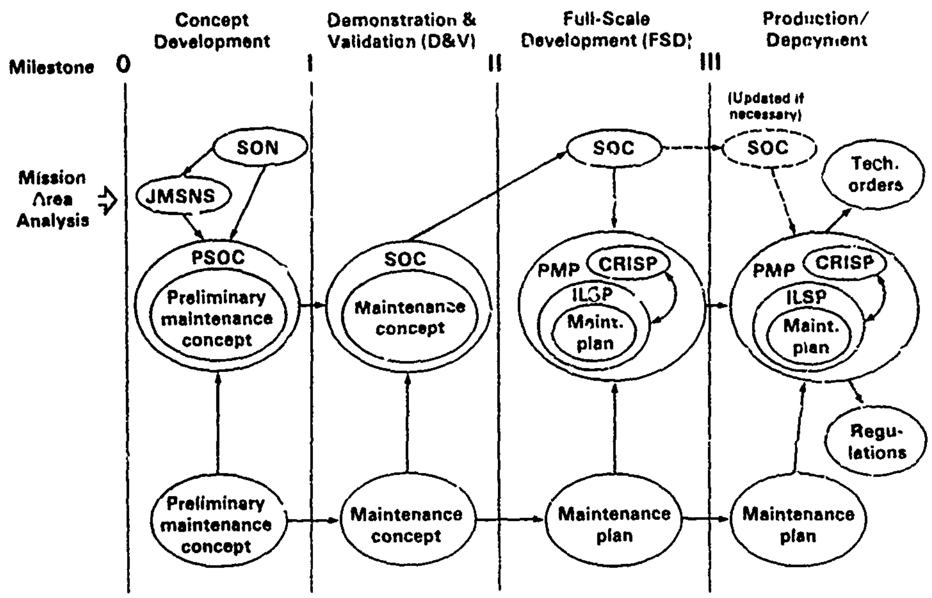
SOURCE: Documents as noted.

Fig. 11—Role of the Baseline Correlation Matrix (BCM)

major original sources of operational requirements documentation is illustrated in Fig. 12.

An unpublished 1982 AFOTEC study that focused on the *PSOC*, however, as well as our review of *SOCs* for several programs, suggested that the timing and content of the *PSOC* and *SOC* limit their utility and prevent them from fulfilling the role of a unified source of requirements. The *SOCs* we examined did not consistently address a full range of operational suitability issues. Air Force personnel also told us that the time required to develop and deliver approved *PSOCs* or *SOCs* limited their impact on RFPs, statements of work, and advanced test planning.[12] A higher priority will have to be attached to the preparation and review of these documents if they are to fulfill the role of a unified source of operational requirements.

[12]The new regulation governing the preparation of operational concepts, *Concepts, System Operational Concepts*, AFR 55-24, forthcoming, is expected to address some of these problem areas. Initiatives at the operating command level will also be needed to address timing and content problems with *(P)SOCs*.



SOURCE: Documents as noted; *Equipment Maintenance, Equipment Maintenance Policies, Objectives, and Responsibilities*, AFR 66-14, 15 November 1978.

Fig. 12—Flow of requirements documentation

OBSERVATIONS

By identifying requirements needing reconciliation in various acquisition documents, the *BCM* elevates the visibility of Air Force requirements expression problems and may also improve the functioning of the acquisition process in the near term. Although the *BCM* does not treat the underlying causes of poor requirements expressions that were outlined earlier in this section, its consistent use may in fact focus management attention on the requirements issue so that causes rather than symptoms begin to be treated, including the greater unification of operational requirements, perhaps using the *SOC* document as a vehicle.

IV. EXPRESSING CONTRACTUAL REQUIREMENTS

Contractors design weapon systems to satisfy contractual specifications, not operational requirements. Contracts must therefore state requirements in such a way as to ensure that the product being developed has the necessary characteristics to permit it to satisfy the operator's needs when working in concert with other Air Force elements. Meaningful contractual specifications, coupled with credible specification compliance measurement techniques, and mechanisms for motivating contractors to comply with specifications are essential features in the overall framework of a sound acquisition strategy.

This section identifies desirable improvements to suitability-related contractual requirements and to means for measuring compliance, identifies some key limitations of past programs for motivating contractors to satisfy requirements,[1] and assesses some of the more important implications, for both the Air Force and its contractors, of new contracting approaches.

We reviewed a mix of old and new programs to gain a historical perspective and to understand future directions in the area. We primarily compared two aircraft programs. The historical perspective was gained by analyzing the initial F-16A/B procurement that began in the mid-1970s. Future directions were assessed by analyzing the C-17 transport aircraft program that entered full-scale development in 1986. We also reviewed some contracting specifications in the F-16C/D follow-on to the original F-16, the Maverick and AMRAAM missiles, and some aircraft subsystems.

CONTRACTUAL REQUIREMENTS MEASURES

Our review of contractual specifications suggested four potentially desirable directions for improvement:

- Make contractors accountable for a broader spectrum of maintenance events

[1]Assessing the comparative utility and applicability of various techniques for motivating contractual compliance is beyond the scope of this study, but other RAND research is addressing this subject, with emphasis on warranties.

- Address a broader set of suitability characteristics with contractual measures
- Increase emphasis on system-level specifications
- Increase emphasis on specifications for production systems

We will discuss each of these while comparing the F-16 and C-17 contracting approaches, and particularly, their treatment of reliability, maintainability, and availability specifications. As an overview, Table 7 shows these specifications for the two programs. The F-16A/B program met a majority of its R&M specifications at the system and subsystem level—as measured by the limited compliance tests used in the program. However, the requirements were not very extensive, and, for the most part, not very stringent.

Table 7

CONTRACTUAL SYSTEM LEVEL R&M
MEASURES IN F-16 AND C-17 PROGRAMS

RELIABILITY

Measure	F-16	C-17
Mission completion success probability	✓	[✓]
Analytical mission reliability		✓
Mean time between maintenance (inherent)	✓	[✓]
Mean time between maintenance (corrective)		[✓]
Mean time between removals		[✓]
Support equipment MTBM(C)		✓

AVAILABILITY

Measure	F-16	C-17
Full mission capable rate	✓*	[✓]
Partial mission capable rate	✓*	[✓]
Maintenance downtime/sortie		[✓]

MAINTAINABILITY

Measure		F-16	C-17
Maintenance manhours per flying hour	on	✓	[✓]
	off	✓	[✓]
	inspections	✓	[✓]
Elapsed time of maintenance tasks (mean/90%)		✓	
Men required per inspection		✓	
BIT	fault detection		[✓]
	fault isolation		[✓]
	false indication		[✓]
Mean manhours to repair			[✓]
Quick turnaround time		✓	✓
maintenance man-minutes			✓
aircrew man-minutes			✓
number of personnel			✓
Mission reconfiguration time		✓	✓
crew size			✓
manhours		✓	
Support equipment MTTR			✓

[ ] Incentivized *Not enforced/reported

SOURCE: *System Specification for Air Combat Fighter System 2185 (F-16A & F-16B)*, Specification Number 16PS001, Code Ident 81755, 16 December 1974; *System Specification for the C-17 Airlift System*, Specification MDC S001, Code Ident 88277, 1 April 1983.

The C-17 program is using more system-level contractual measures, although they represent an overprescribed set in the sense that some reliability measures (e.g., mean time between removals and mean time between maintenance (inherent)), are at least partially subsets of other measures (e.g., mean time between maintenance (corrective)). In comparing contractual requirements, the absolute numbers of measures used are less important than the scope and significance of the individual measures. That is where the two program approaches differ appreciably and where we will focus our attention.

System Reliability

Specifications of reliability for most types of equipment have typically covered only a fraction of system or subsystem demands for maintenance. This is a consequence of (1) narrowly defined measures that exclude several important categories of resource-consuming maintenance events, and (2) contractual compliance test ground rules that further narrow the list of contractually chargeable failures.

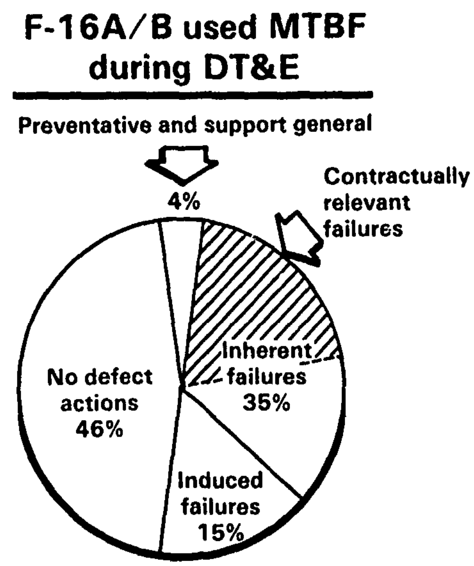
Inherent failures are those failures caused by an item's own internal failure patterns. Such failures are usually confirmed at intermediate or depot levels of maintenance. *Mean Time Between Failures (MTBF)* is the most commonly used measure to express inherent reliability in reliability specifications. For each inherent failure maintenance event on a fighter aircraft, there are typically two to three times as many other kinds of maintenance events categorized as induced failures or no-defect events. Despite the frequency with which they occur, these latter kinds of maintenance events have commonly been excluded from contractual reliability specifications.

Induced failures describe situations when equipment cannot meet its minimum specified performance requirement due to some induced condition and not due to its own internal failure pattern. Induced failures include such events as damaging a panel fastener or electrical connector when accessing a suspected faulty Line Replaceable Unit (LRU), foreign object damage to an engine, damage due to malfunctioning of associated equipment, operator errors, etc.

No-defect events are more common and include situations when maintenance resources are expended due to policy, modification, location, or cannibalization without any actual malfunctioning of an item. No-defect events include such things as removing one unit to facilitate maintenance on another, swapping LRUs in and out of the aircraft when troubleshooting, removing an LRU thought to be malfunctioning that subsequently is shown not to be, and removing a good LRU from one aircraft to replace a malfunctioning LRU on another aircraft (cannibalization).[2]

Induced failures and no-defect maintenance events consume appreciable maintenance man-hours,[3] tie up support equipment, add to spares requirements, and keep airplanes out of commission just as inherent failures do. The F-16A/B system-level specification for reliability, Mean Time Between Inherent Failures (MTBF), excluded induced and no-defect maintenance actions, which accounted for two-thirds (65 percent) of all maintenance actions during Development Test and Evaluation (DT&E) (see Fig. 13).

DT&E ground rules excluded other actions falling into any of 15 so-called nonrelevant hardware failure categories, such that only about 19 percent of all maintenance actions during DT&E were considered



SOURCE: DiGiovanna and Eischens (1979).

Fig. 13—Contractually relevant failures in the F-16 program

[2]H. D. Rue and R. O. Lorenz, *Study of the Causes of Unnecessary Removals of Avionic Equipment*, Rome Air Development Center, RADC-TR-83-2, January 1983.

[3]Maintenance data for the F-16A worldwide during calendar year 1982 indicate that maintenance man-hours expended dealing with induced and no-defect failures accounted for 37 percent of all flight-line man-hours; roughly another third was expended coping with inherent failures (30 percent); the remainder (33 percent) of the man-hours were expended accomplishing preventative and support general activities (preventative maintenance and servicing the aircraft).

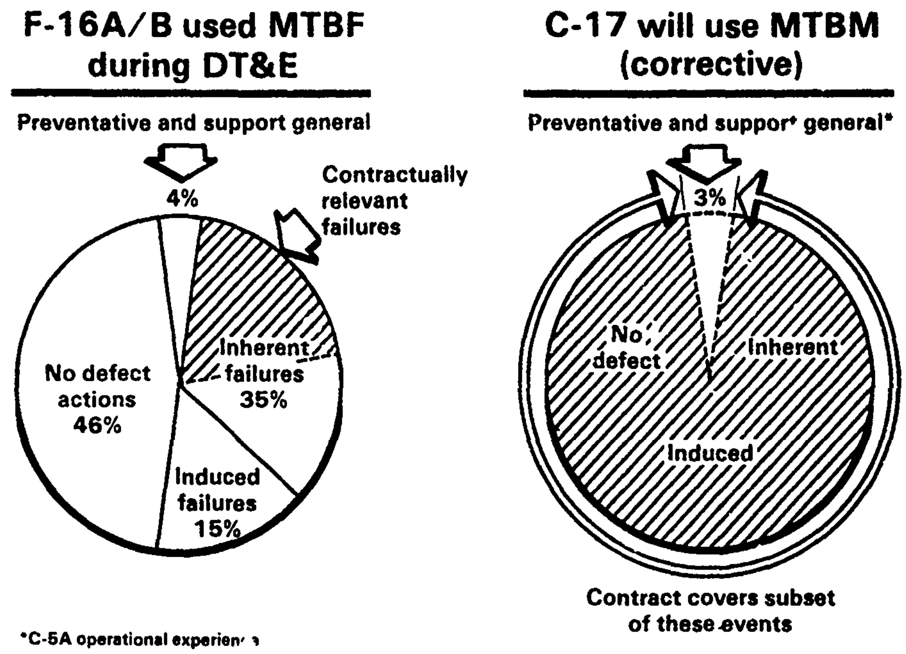
contractually relevant and chargeable against the F-16 system reliability specification. Under this restrictive definition of reliability, the F-16 met its system-level reliability requirement during a 97-flight-hour reliability demonstration period.

Given their effect on narrowing contractual accountability, ground rules for categorizing failures must receive the same scrutiny as the definitions of specification measures to put the meaning of compliance testing results in proper perspective. In F-16 DT&E testing, ground rules excluded failure events such as the wearout of life-limited items like tires when their life had been exceeded, software reliability problems, and those failures identical to previous failures for which a corrective Engineering Change Proposal (ECP) had been approved but not implemented on the particular aircraft undergoing test.[4] The latter exclusion, perhaps the most significant of the 15 failure categories, has been common practice in test programs, because there is often inadequate time to go through a full cycle of problem identification, correction, and retest during a given test phase. It is argued that excluding failures for which apparent fixes have been identified gives a sense of the potential improvement in reliability that might be expected once those fixes are indeed incorporated. However, such an approach obviously also carries with it some element of risk that corrective maintenance actions may subsequently be shown to be ineffective.

The use of narrow contractual measures of reliability has several potentially undesirable effects. First, contractors tend to focus their design efforts on those aspects of system reliability on which they will be evaluated, i.e., inherent reliability, rather than taking a more balanced approach that recognizes the influence of maintainability characteristics on reliability (reflected by the frequency of induced and no-defect maintenance events). Second, because a measure such as MTBF is not a meaningful term in an operational sense, there is a risk that satisfaction of a narrowly defined contractual requirement may not be synonymous with delivery of an operationally suitable system, and, in fact, may lead to different assessments of a system's adequacy by the developer and the operational test community.

Several recent programs are expanding contractual accountability for reliability through the use of broader measures. The C-17 program specification for contractual system reliability encompasses not only inherent failures, but also induced and no-defect maintenance events (see Fig. 14). Although we cannot be sure of the exact fraction of maintenance events that will be contractually chargeable until detailed

[4]For the full list of failure event exclusions, refer to DiGiovanna and Eischens (1979), App. E and p. 19.



SOURCE: *System Specification for the C-17 Airlift System*, Specification MDC S001, Code Ident 88277, 1 April 1983; DiGiovanna and Eischens (1979).

Fig. 14—Comparison of contractual system reliability in F-16 and C-17 programs: distribution of failure causes

test ground rules are written, the degree of specificity in the system specification suggests that the coverage will be much greater than that which has been typical in past aircraft programs. Hence, the scope of the definition of the reliability measure will have the beneficial effect of requiring that the contractor actively monitor and control failures in each of the three major categories—inherent, induced, and no-defect.

Under terms of the reliability improvement warranty for the LN-39 standard navigation unit installed in the A-10 and the F-16, units removed because of an indicated failure that subsequently retest OK (RTOK) are counted as failures when they exceed 5 percent of total failures. Because of the broadness of definition of the reliability measure and RTOK rates that have been in excess of 5 percent, the contractor has devoted considerable design attention to the Built-in Test (BIT) software to bring down the RTOK rate. Hence, the broader measure of contractor performance, coupled with field data collection of the equipment's performance to measure compliance, has had the desired effect of focusing the design improvement process on an operationally relevant problem.[5]

Mission Reliability

Mission reliability describes the probability of a system completing an attempted mission successfully, which depends on both the reliability of the hardware and the redundancy built into the system. Sparse definitions of mission reliability and the ground rules for measuring compliance during testing have limited the meaning of this contractual specification.

The F-16 specification never actually defined the term "mission reliability"; instead, it listed equipment that had to operate properly for mission success.[6] This sparse definition gave the F-16 Joint Test Force great latitude in defining the terms for measuring mission success. DT&E ground rules categorized as *nonrelevant* 19 kinds of mission failure events causing aborts.[7] As a consequence, about half of all ground and air aborts were excluded from the calculation of mission reliability. The ground rule that excluded those failures for which corrective ECPs had been identified but not incorporated appears to account for the majority of exclusions.[8]

Other F-16 ground rules for excluding mission failures seem unnecessarily restrictive, although test documentation is not complete enough to gauge the individual impact of particular ground rules on contractually measured mission reliability. For example, failures discovered prior to engine start initiation that resulted in an aircrew decision to ground abort were excluded. This measurement approach excludes an important aircraft attribute, which is its dependability in flying a scheduled mission whose planned takeoff time is set well before the aircrew climbs into the cockpit and attempts to start the engine. Another ground rule essentially excluded intermittent mission failures that erode operator confidence in an aircraft. Another excluded those failures that occurred due to equipment which did not meet specification, drawing, or other conformance requirements. This

[5] *Standard Navigation Unit, Reliability Improvement Warranty*, manufacturer's briefing, 13 May 1985.

[6] *System Specification for Air Combat Fighter System 2185 (F-16A & F-16B)*, Specification Number 16PS001, Code Ident 81755, 16 December 1974, p. 8, App., Sec. 40.

[7] DiGiovanna and Eischens, App. E.

[8] For example, there were repeated failures of the jet fuel starter after apparent fixes had been identified but not incorporated because of the constrained time schedule of testing, resulting in numerous ground aborts that were not contractually chargeable.

seemingly is a contractor responsibility. Failures due to foreign object damage (FOD) were also excluded. This failure category seems to deserve greater disaggregation, since at least some FOD incidents can be traced to poor air vehicle design.

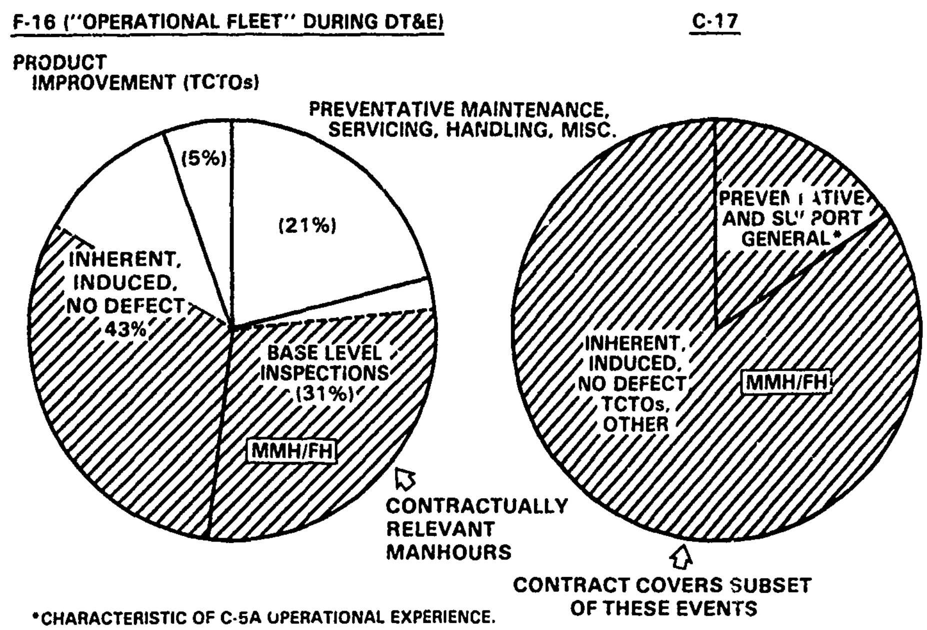
Detailed test ground rules have not been written for the C-17, so we cannot make a direct comparison, although the specification definition of mission reliability is much more comprehensive. It defines mission reliability as the probability that a scheduled mission will be completed without experiencing an on-equipment failure or performance degradation that would result in an air abort, ground abort, or mission deviation.[9] The more comprehensive definition of mission reliability in the C-17 specification directly addresses the dependability in scheduling attribute not captured in the F-16 program. Other comparisons must await the formulation of test ground rules some years hence, but it is apparent that opportunities exist for making contractors more accountable for mission reliability.

Maintainability

Our review of contractual measures of aircraft maintainability such as Maintenance Man-hours per Flying Hour (MMH/FH) suggests they have been more broadly defined than reliability, although the contractor is still not accountable for an appreciable fraction of the maintenance effort. Contractual MMH/FH in the F-16 contracting arrangement encompassed about 60 percent of operationally defined MMH/FH (see Fig. 15). About half of the excluded man-hours (21 percent of the total) were those expended in so-called Support General tasks such as servicing, handling, washing, cleaning, corrosion prevention treatment, arming, disarming, etc.[10] Another 12 percent of the total man-hours were exclusions from the corrective maintenance category, such as excluding man-hours expended to correct malfunctions for which an ECP had been approved but not yet implemented on

[9]Air aborts describe missions in which on-equipment failures or performance degradations result in unscheduled landings. Ground aborts describe missions in which on-equipment failures or performance degradations prevent takeoff within a prescribed time of the scheduled takeoff time and which are discovered after the aircrew arrives at the aircraft. Mission deviations include failures or degradations other than air or ground aborts that prevent the aircraft from completing all mission objectives, takeoff delays in excess of a prescribed time for maintenance after the first takeoff of the mission, and changes in the aircraft tail number scheduled for a particular mission within a prescribed time of scheduled takeoff time. *System Specification for the C-17 Airlift System*, Specification MDC S001, Code Ident 88277, 1 April 1983, pp. 60-9, -10, -11.

[10]These excluded man-hours are described by Air Force Work Unit Codes (WUCs) 1,2,5,6,7,8, and 9. The look phases of scheduled and special inspections, WUCs 3 and 4, were included in the man-hour accounting.



SOURCE: *System Specification for the C-17 Airlift System*, Specification MDC S001, Code Ident 88277, 1 April 1983; DiGiovanna and Eischens (1979).

Fig. 15—Comparison of contractual system maintainability in F-16 and C-17 programs: distribution of maintenance man-hours per flying hour

the aircraft undergoing test. Although detailed test ground rules will likely exclude some operationally defined man-hours from the contractual accounting, the C-17 program's maintainability measure includes all major categories of maintenance man-hour expenditures, fulfilling the need for broader contractor accountability for important aspects of R&M.

The C-17 contract also specifies another important element of maintainability, diagnostic system performance, quite differently from previous Air Force programs. Rather than prescribing a required usage of built-in test (BIT), the contract specification leaves the mix of BIT, test equipment, and manual diagnostic methods up to the contractor to decide, with the overriding stipulation that working in combination, they must be able to detect and isolate all faults on the aircraft and permit satisfaction of the other maintainability requirements measured in terms of man-hours and elapsed time to accomplish specific tasks. When and where the contractor elects to use BIT, he must meet requirements for fault detection, fault isolation, and false-indication performance. This approach leaves the detailed design decision up to the contractor while still stipulating the overall diagnostic performance the Air Force expects.[11]

Availability

Defined most simply, availability is the ratio of system uptime to total time. It translates reliability, maintainability, and logistics supportability characteristics into a measure of the fraction of time the system is available to the operator. Availability is very difficult to specify in a contractual sense because the test environment, support system, and utilization rate are usually not representative of that which would be experienced in field operations, and hence direct measurements during compliance testing are not very reliable.

Some programs have specified an *inherent availability* that considers only the influence of inherent reliability and the mean time to repair inherent failures, but this type of definition suffers from the same lack of operational meaning as system reliability specifications that encompass only inherent failures. A somewhat broader measure of availability is termed *achieved availability*, which considers all active corrective and preventative maintenance time, but which also excludes various delays waiting for maintenance or supply resources, or other miscellaneous delays, which can be consequential in an actual operating environment.[12]

A third approach measures availability directly, but simulates various supply delays in controlled testing by applying standard minimum delay times for supplying a replacement item to restore an aircraft to mission-capable status and for replenishing spares stocks.

Our review identified an availability specification for the F-16, but we found no record of its measurement in F-16 DT&E test reports. In contrast, the C-17 program includes a comprehensive contractual specification for availability to be measured during development testing, during a special controlled test phase in which supply delays will

[11]This approach is consistent with recommendations made by AFOTEC based on their experiences in testing automated diagnostic systems. *Automated Diagnostic Systems*, undated AFOTEC briefing. It does, however, require care in the formulation of requirements to ensure that other design goals are not compromised. For example, the absence of constraints on manpower (or maintenance effort) or on the quantity of support equipment, could conceivably cause a contractor to rely more heavily than the user would desire on manual diagnostic techniques or on the use of test equipment, either of which could compromise mobility or other acquisition objectives.

[12]For a thorough discussion about approaches for specifying and measuring availability, see *Logistics Assessment*, Air Force Operational Test and Evaluation Center Pamphlet 400-1, 29 February 1984, Chap. 16.

be simulated, and during early operations. To the authors' knowledge, this represents the most ambitious attempt to contractually specify and measure availability. The C-17 prime contractor regards availability as perhaps the most difficult of all the program's contractual specifications to measure in a realistic and fair manner, and hence believes there is more risk associated with attaining the availability specification than with some of the other parameters that can be measured in a more straightforward manner. The C-17 program experience should help establish the feasibility of this approach for contractually specifying and measuring availability.

Specifying Other Important Suitability Characteristics

Two other generic areas of suitability characteristics, system attributes that influence mobility and resilience to attack, have historically received limited treatment in contractual specifications. Section II has already discussed important considerations in expressing mobility needs. In addition to functional performance considerations that have traditionally been contractually specified (e.g., ferry range), other critical considerations include transportation resource requirements for deployment and the responsiveness of a system and its support elements. Functional performance attributes can be measured directly, while the latter two characteristics will have to be demonstrated through a combination of actual demonstrations and analytic justification. For example, testing should help support computation of personnel, War Reserve Spares Kit, and ground support equipment needs. The weight and volume characteristics of these can then be used in loading models to determine transportation resource requirements. Elements of responsiveness can be measured in field demonstrations of the time to break down for transport, and subsequently set up for operation, various support elements. Without minimizing the difficulty of writing a mobility specification, there seem to be opportunities to do more in this area than has been done in the past.

Resilience to attack is a complex function of many factors, including threat attack scenarios, basing and support environments, operational procedures, and weapon system features. Although no inclusive contractual measure for this attribute seems possible, it does seem feasible to specify certain weapon system features that will contribute to resilience to attack. For example, in measuring aircraft battle damage repair capability, one could specify a contractual requirement to build representative test sections of aircraft structure through which prescribed weapons would be fired, and that would then have to be repaired in a prescribed period of time using field repair techniques and then be tested for strength. Times to accomplish critical maintenance tasks when wearing chemical ensembles could be specified. Other system characteristics already discussed and subject to specification (such as reliability, maintainability, mobility, etc.) also contribute to the specification of resilience to attack both directly and indirectly, through their influence on basing and support needs and operational procedures.

Other Considerations in Specifying Requirements

In comparing F-16 and C-17 R&M contract specification approaches, we identified two other potentially important differences: the maturity of the equipment to which the specifications applied and the reliance on system versus subsystem specifications.

All *system-level* contractual specifications for R&M of the F-16 were written only for the pre-production system. The system specification included *goals* that the Air Force wanted to achieve in subsequent production aircraft two years after delivery of the 24th aircraft to TAC, but these goals placed no contractual obligation on the supplier.[13] In contrast, the C-17 program includes extensive system-level specifications for R&M and availability of the production system that will have to be demonstrated by adherence to contractually specified growth curves calling for continual improvement as flight hours are accumulated during testing and during field operations to 6 months after declaration of Initial Operational Capability (IOC).

If pre-production system requirements and production system goals differed by only a modest amount, then one might argue that incorporation of the corrective engineering changes identified during testing of the pre-production system plus routine quality assurance during manufacturing would provide assurance that goals for the production system would ultimately be satisfied, but several factors undermine this argument. First, pre-production requirements and mature production goals can differ substantially, as the comparison of F-16 requirements and goals shown in Table 8 illustrates. In fact, if one uses the availability figures in the table to calculate mission *unavailability*, we can see that it must be halved (30 to 15 percent). Similarly, mission *unreliability* must be cut by one-third (15 to 10 percent). Significant improvements in hardware reliability and maintainability are also

---

[13]Lab demonstrations of the reliability of the following production items were required: radar, radar/electro-optical display, stores management set, head up display, fire control computer, flight control computer, attitude director indicator. Nine avionics items were also covered by a repair cost warranty, and the radar transmitter and HUD electronics were also covered by MTBF guarantees.

Table 8

F-16A/B CONTRACTUAL REQUIREMENTS AND GOALS

Parameter	Pre-production Requirement	Mature Production System Goal	Change (Percent)
Availability	70%	85%	21
Hardware reliability (MTBF)	1.75 hrs	2.9 hrs	66
Mission reliability	85%	90%	6
Maintainability	20 MMH/FH[a]	12 MMH/FH	-40

SOURCE: *System Specification for Air Combat Fighter System 2185 (F-16A & F-16B)*, Specification Number 16PS001, Code Ident 81755, 16 December 1974.

[a]MMH/FH is maintenance man-hours per flying hour.

called for. Second, compliance testing of pre-production systems generates a reservoir of unimplemented and untested hardware and software design changes, introducing considerable uncertainty about the R&M and availability characteristics that the ensuing production system will ultimately exhibit. Third, manufacturing methods can change when moving from the pre-production to the production system, introducing additional uncertainties about characteristics of the production system.

Given the availability of sufficient time and resources for compliance testing, specifying system-level requirements for the production system would seemingly provide more rigorous contractor accountability for achieving the level of performance ultimately deemed necessary to provide a system that can satisfy operational requirements. Requirements specified for the production system would not necessarily be the same as those expected of the mature system, but rather would be intermediate values lying between pre-production and mature values, as charted by growth curves.

Another potentially important difference between F-16 and C-17 programs involves the manner in which subsystem specifications are prescribed. In the C-17 program, this authority is delegated to the prime contractor, who allocates subsystem R&M requirements in contracts with its vendors, consistent with the need to satisfy binding system-level requirements for which it is responsible to the Air Force (see Table 7). In the F-16 program, the Air Force specified contractual requirements to be demonstrated in the laboratory and/or the field for selected subsystems, items, and electrical/mechanical equipment.[14] In some cases the Air Force contracted directly with the subcontractor for the item, such as the engine, whereas in other cases the prime contractor had responsibility for ensuring that the Air Force specification was met.

C-17 program officials have cited two principal advantages of the system-level requirements approach. First, it gives the prime contractor the freedom to make design trades, and second, it forces compatibility between system design and the design of support equipment and maintenance concepts. The reliability, maintainability, and availability of the support equipment influence those characteristics of the aircraft system as well, and hence the support equipment must be designed in such a way that system-level requirements for R&M and availability can be met.

The system-level parameters selected were limited to those that affected aircraft availability to the using command and the operational readiness of the aircraft system, which are more important considerations to the operator than the reliability or maintainability of any individual subsystem, item, or component.[15] The delegation of subsystem requirements responsibility is therefore a means for giving the contractor greater operating flexibility to satisfy the broader than normal system-level requirements for R&M and availability. This approach must compete directly with the historical inclination of program offices to exert tight control over all aspects of system development.[16]

Other Systems

Although our focus in reviewing contractual requirements approaches was on original aircraft development programs, we briefly examined the F-16C/D upgrade to the F-16A/B and the Maverick and AMRAAM missile programs.

---

[14]The flight control system is an example of a subsystem for which a reliability requirement was specified. The flight control computer is an example of an item, while the flight control panel is an example of electrical/mechanical equipment.

[15]See Emmelhainz (1983).

[16]In practice, although not spelled out directly in the specification, there was also some recognition of the desirability of the system-level perspective in the F-16 program. The overall air vehicle pre-production specification calling for a 1.75 hour MTBF was regarded as the key reliability requirement in the program, and the failure of a subsystem to meet its requirement did t necessarily trigger a contractually mandated corrective action, so long as the system level requirement was met.

The F-16C/D is an upgrade to the F-16A/B incorporating substantial improvements to the fire control system (including the new APG-68 radar) and specialized weapon delivery systems (including incorporating the ability to use the AMRAAM missile and the LANTIRN system). The rest of the airplane is substantially the same as the A/B version.

A review of the contractual specifications of the F-16C/D shows no appreciable expansion in the scope of R&M requirements in terms of either the measures used or the compliance testing called for. There are no system- or subsystem-level specifications for development articles, although goals are stated for the mature production system and major subsystems. Many of the measures used are the same, including MTBF, mission reliability, and maintenance man-hours per flying hour. There are reliability specifications for some development and production items, such as the fire control computer and multifunction display set. All contractual compliance for these items was to be demonstrated in the laboratory.[17]

The lack of system-level R&M specifications for the model change from the A/B to C/D progression may reflect the fact that a substantial part of the aircraft remains the same, and the model change development focus is not on improving the overall reliability of the system. The lack of R&M specifications for the performance of new items when operating in the airplane reflects the traditional preference for measuring the specification compliance of items in the controlled environment of a laboratory, and perhaps also schedule and fiscal constraints on the nature and extent of field compliance testing.[18] Reliance on laboratory compliance testing to demonstrate satisfactory performance obviously carries with it more risks, since such testing, for a variety of reasons, has usually been an imprecise indicator of performance to be expected in the airplane during operations.[19]

Appreciable progress is apparent in the specification of contractual requirements for newer Air Force tactical missiles. The AMRAAM program is the most prominent example of the high visibility accorded R&M and availability in contractual specifications, testing, and incentive arrangements. It includes a comprehensive set of reliability specifications for all phases of the missile life cycle, including the dormant storage phase, the uploaded but non-powered, non-airborne phase, the captive carry flight phase, and the launch to intercept phase. Both

[17]R&M specifications for the radar were not available at the time of our review, so these generalizations do not necessarily apply to that item.

[18]The F-16C/D has, however, undergone developmental and operational field testing.

[19]Kern and Drnas, (1976).

field and environmental chamber compliance testing is being conducted, including a storage reliability test program with incentives.[20] And, in contrast to many aircraft programs, production versions of tactical missiles, including AMRAAM, are usually subjected to contractual compliance testing both in new development programs and to qualify missiles produced by second sources.

Summary Observations About Contractual Requirements Measures

Our review of contractual requirements suggested several potentially desirable improvements in the specification of suitability-related system characteristics. Some newer Air Force programs are incorporating a number of these improvements, and this experience should prove valuable in identifying the most effective ways of implementing the improvements in actual programs. In the case of the C-17, it will be the 1990s before the effectiveness of the initiatives can be fully assessed.

Our review suggests the need for broader contractual accountability for those events that influence a system's demands for maintenance. This will require the use of contractual measures encompassing more categories of maintenance events and the use of compliance test ground rules that do not compromise those measures.

It appears desirable to use contractual measures that address suitability characteristics critical to resilience to attack and mobility. Although all-encompassing contractual measures for these characteristics do not seem apparent, there are opportunities to improve significantly the specification of system characteristics that contribute to these two aspects of operational suitability.

More contractual specifications at the system level appear desirable to focus design attention on how the subsystems work as an integrated unit. This will become even more important in the future as fire control, propulsion, flight control, and navigation functions are increasingly integrated.

Greater use of system level contractual specifications and associated compliance measurements for production systems is desirable. Extrapolating performance exhibited by pre-production systems and stating only goals for production systems involve much uncertainty.

---

[20]One missile program we reviewed included dormant reliability specifications, but for reasons of cost, the contract did not call for the contractor to demonstrate satisfaction of the specification.

COMPLIANCE TESTING

For more stringent R&M specifications to have credibility, they should be accompanied by more meaningful contractual compliance testing than has been the case in the past. This testing can take place both in a laboratory setting for subsystems and at the field level for systems and subsystems. To identify potentially desirable improvements in compliance testing, we compared the F-16 field testing experience with that planned for the C-17.[21]

F-16 Compliance Testing in the Field

Several aspects of F-16 compliance testing limited its meaning, including a small sample of aircraft, limited calendar time for testing, a limited accumulation of flight hours, and a nonrepresentative support environment, as well as narrowly defined contractual measures and test ground rules for contractual relevance. The acquisition setting within which the F-16 was developed, including extreme schedule pressure with considerable concurrency of development and production, constrained the quality of compliance testing possible, but that testing also reflects a limited focus on contractual R&M measurements that the Air Force is currently trying to overcome through policy pronouncements and changes in acquisition procedures.[22]

The calendar time, flight hours, and test articles used to measure F-16 contractual compliance were only a fraction of those associated with overall Development Test and Evaluation (DT&E). DT&E began 23 months after the start of full-scale development and lasted for two years using 8 DT&E aircraft and 3 production aircraft. R&M data were collected during combined development and operational testing, with the total flight hours accumulated for contractual and operational evaluations shown in Table 9.[23] Contractual R&M values were computed throughout the test program; however, they were used only "as a

[21]This report does not address various approaches to component, item, or subsystem level testing in the laboratory or manufacturing environment, a subject that has been addressed in numerous reports and papers over the past decade. Environmental stress screening, burn-in, environmental chamber testing, and other approaches are important means for improving the quality of the product that is ultimately subjected to system-level testing.

[22]For a discussion of some of the factors that shaped the F-16 schedule, see Rich, Stanley, Birkler, and Hesse (1981). Section V of this report looks beyond mere compliance testing in discussing how the F-16's acquisition setting influenced both developmental and operational testing.

[23]Most sorties supported both contractual and operational measurcment objectives. Some dedicated sorties were flown and that accounts for the differences in flight-hour totals between the contractual and operational fleet shown in Table 9.

Table 9

F-16 USE DURING DT&E

Fleet	Number of Aircraft	Calendar Months of Testing	Aircraft Months of Testing	Flying Hours	Sorties
Operational	8 DT&E/3 prod	25	108	2184	1707
Operational (prod acft only)	3 prod	5	8	132	119
Contractual	8 DT&E/3 prod	25	108	1900	1591
Contractual compliance measurement	2 DT&E	3	6	97	80

SOURCE: DiGiovanna and Eischens.

gauge indicating how the contractor was doing."[24] Formal contractual compliance with R&M specifications was measured during the last 3 months of development testing using 2 DT&E aircraft that accumulated a comparatively limited 97 flight hours.

The use of only 2 aircraft and the limited time base of compliance measurements raise questions about whether the compliance measurements obtained were representative of the actual R&M characteristics of the system at that time. R&M performance can vary considerably from one aircraft to another, whether in a test or an operational environment. The two DT&E aircraft involved in R&M compliance testing exhibited such variability. They flew almost identical numbers of missions and yet one experienced 17 aborts whereas the other experienced only a single abort.[25] Which aircraft, if either, accurately reflected the mission reliability delivered by the contractor?

The considerable number of subsystems (8 of a total of 23)[26] having predicted reliabilities (mean time between failures) greater than the time-base for compliance testing—97 flying hours—also casts doubt on

[24]DiGiovanna and Eischens (1979).

[25]Similar inconsistencies in aborts from aircraft to aircraft were exhibited by the three production aircraft involved in the test program. One experienced 15 aborts, another 9, and a third 3.

[26]Those eight subsystems were the crew station (105-hour mean time between failures predicted by the contractor), auxiliary power (113 hours), miscellaneous utilities (590 hours), service life monitoring (1050 hours), VHF communications (210 hours), interphone (420 hours), air-to-ground IFF (145 hours), and ECM system (125 hours).

the representativeness of compliance test results. During compliance testing, nine subsystems experienced no contractually relevant failures, i.e., using the contractually defined measure for reliability—MTBF, and the restrictive test ground rules described earlier—none of the kinds of failures these subsystems experienced were counted as being contractually chargeable. Five of those subsystems had predicted MTBFs greater than 97 hours. With testing of such limited duration, one cannot hope to get a reliable indication of a system's actual reliability characteristics.

A further illustration of uncertainty about the representativeness of the compliance measurements is the disparity in the reliability observed during the brief contractual compliance measurement period and the more extended "contractual fleet" measurement of R&M shown in Table 10. The system hardware reliability measured during the brief contractual measurement period was 67 percent greater than that observed over the 1900 flying hours accumulated by the contractual fleet.[27] Differences in other characteristics were not as large, but still appreciable.

Table 10

COMPARISON OF R&M MEASUREMENTS

Fleet	System Reliability (MTBF-hr)	Mission Reliability (%)	Maintainability (MMH/FH) (Corrective)	(MTTR) (hr)
Contractual fleet (1900 flying hours)	1.5	90.8	9.6	2.6
Contractual compliance measurement (97 flying hours)	2.5	85.7	7.5	2.0

SOURCE: DiGiovanna and Eischens.

NOTES: MTBF is mean time between failures; MMH/FH is maintenance man-hours per flying hour; MTTR is mean time to repair. Mission reliability is measured by the percentage of attempted missions completed successfully.

[27]Differences in reliability measurements of many individual subsystems were much greater. The apparently conflicting result in Table 10 of a higher system reliability having a lower mission reliability is most likely the product of two factors. First, because of redundancy in some functions, a mission can succeed even when certain equipment fails. Hence, for a given mission, the system reliability category might log a failure even when the mission was recorded as a success. Second, differences in test ground rules for counting system failures and mission failures can cause apparent inconsistencies. For example, failures discovered before engine start were not counted as mission failures, whereas

Abbreviated compliance testing has implications for both the buyer (the Air Force) and the producer (the contractor) of the equipment. The Air Force runs the risk of accepting equipment that apparently meets reliability specifications when in fact its true reliability is less than the specification. The contractor faces the opposite risk: During abbreviated testing, the equipment may fail to meet specifications when in fact the true reliability exceeds specifications.[28]

Other factors limiting the meaning of F-16 compliance measurements included a lack of most intermediate level test equipment, including the Avionics Intermediate Shop (AIS)[29] —a critical set of test equipment needed to perform off-aircraft avionics repair—contractor control and maintenance of some aircraft, a lack of adequate spares, and an inadequate system for tracking the status of the airplane's configuration.

The lack of most intermediate level test equipment precluded much evaluation of the off-equipment maintenance characteristics of the F-16, since most maintenance that would normally have been performed at the intermediate level was accomplished at the prime contractor's facility or at those of its vendors. The equipment, personnel, and procedures used for repairs at contractor facilities were not representative of the field environment.

The contractor maintained and controlled several aircraft during DT&E, some for the entire test program and others for part of the program, using nonstandard maintenance procedures and policies. These aircraft accounted for about 73 percent of the flying hours and organizational man-hours, although test documentation does not indicate the extent of contractor involvement in maintenance and flight of aircraft used for compliance measurements.

A lack of spares, a situation common to many test programs, limits or precludes accurate measurement of system availability.

Not having a good configuration status system made it difficult to evaluate the effectiveness of aircraft improvements because the "pedigree" of some of the equipment on aircraft was not known. Seemingly, this shortcoming would also undermine efforts to apply one of the key test ground rules discussed in the previous subsection, that being the

---

they could be counted against system reliability.

[28]There are well-developed methodologies for considering such risks during test design, although one must make some rather restrictive assumptions about the way systems fail in order to compute the risks conveniently. See *Reliability Design Qualification and Production Acceptance Tests: Exponential Distribution*, MIL-STD-781C; and Zwanziger (1985).

[29]Section V discusses how changes in subsystem development strategies can facilitate the earlier availability of subsystems and their associated support equipment for testing.

classification of failures as contractually chargeable or not chargeable according to whether fixes had already been identified but not implemented on particular aircraft.

The aforementioned factors severely limited the meaning of F-16 system-level R&M compliance measurements made during the last three months of DT&E. The DT&E test report acknowledges the limitations of the test and appropriately urges caution in interpreting the results. If the use of more stringent R&M specifications is to have any appreciable effect, future programs must strive to avoid the limitations that characterized F-16 compliance testing.

C-17 Compliance Testing

Our review of the C-17 program indicates a more rigorous approach to measurement of contractual compliance with R&M and availability specifications. Table 11 illustrates some of the key differences in the F-16 and C-17 arrangements.

In contrast to fighter aircraft programs, the C-17 contract calls for only one DT&E test aircraft, with the remainder of the aircraft used

Table 11

COMPARISON OF SYSTEM-LEVEL COMPLIANCE MEASUREMENT IN F-16 AND C-17 PROGRAMS

Aspect of Test	F-16A/B	C-17 (Projected)
Aircraft sample	2 DT&E	More than 12 prod acft
Time period	3 months	2.5 years
Flying hours	97	10,000
Sortie emphasis	Developmental/operational	Operational
Test environment setting	Test base	Test/operational base
Measurement relative to Milestone III	14 months after	Before and after
End of compliance measurement	End of DT&E	6 months after IOC

SOURCE: DiGiovanna and Eischens (1979); *System Specification for C-17 Airlift System*, Specification MDC S001, Code Ident 88277, 1 April 1983; C-17 System Program Office, October 1983; *C-17 Preliminary Design Review (PDR), Reliability and Maintainability*, Douglas Aircraft Company, McDonnell Douglas Corporation, 7-17 May 1985.

for testing and field operations being designated production aircraft. This is probably impractical for fighter aircraft programs, which normally face greater technological challenges.

Instead of measuring compliance with system level specifications only during formal development testing, the C-17 program will measure compliance during FOT&E and for six months of operations as well. This accounts for the relatively large sample of aircraft participating in compliance measurements, the long calendar time, the large accumulation of flying hours expected by the conclusion of compliance measurements, and the operational environment for much of the compliance measurement period. All of these factors should enhance the compliance measurements.

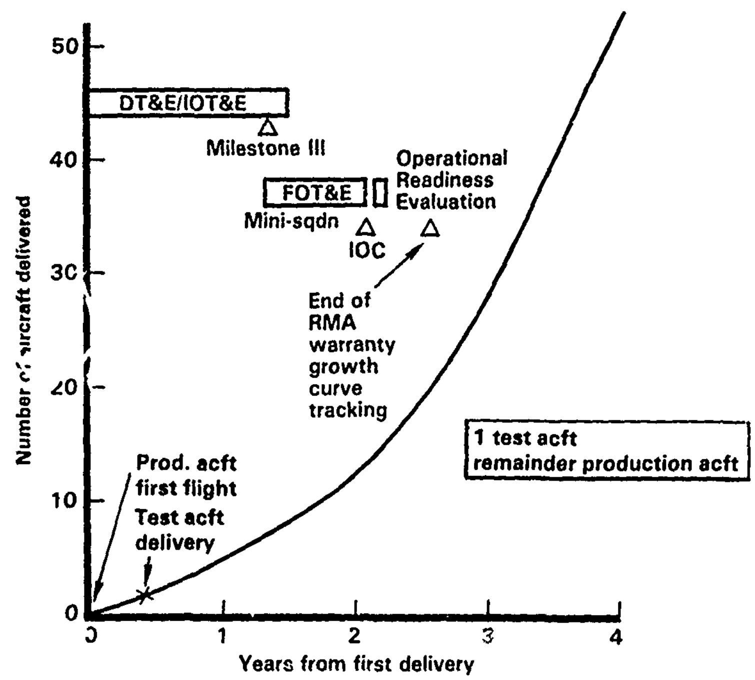
Rather than having to satisfy point estimate R&M requirements in one snapshot evaluation, the contractor will have to demonstrate continuing improvement in R&M and availability parameters. Using established growth curves for R&M and availability from the contract,[30] performance will be assessed at a 50-flying-hour check point during DT&E and at the conclusion of DT&E/IOT&E. (Figure 16 shows the projected schedule for testing and aircraft deliveries.) Then compliance will be measured on a continuing basis through growth-curve tracking that will begin during FOT&E at an operational base. Under terms of the contract, the contractor is obligated to take corrective action when performance fails to meet growth-curve thresholds. After FOT&E, a special intensive one-month structured Operational Readiness Evaluation will determine the award of an incentive fee for exceeding R&M and availability thresholds. Contractual compliance through growth-curve tracking, and the contractual obligation for taking actions necessary to comply with those curves, continues for 6 months after declaration of Initial Operational Capability.

Desirable Attributes for Compliance Testing

Every program has its own specific needs and requirements. Nonetheless, several attributes of the C-17 compliance measurement approach seem very attractive and worthy of consideration for use in future programs, even though implementation details may differ.

More calendar time for making compliance measurements, during both formal testing and operations, should enhance opportunities for incorporating fixes and measuring their effect on contractual compliance directly, rather than having to artificially adjust the number of

[30]*System Specification for C-17 Airlift System*, Specification MDC S001, Code Ident 88277, 1 April 1983.



SOURCE: C-17 System Program Office, October 1983; *C-17 Preliminary Design Review (PDR), Reliability and Maintainability*, Douglas Aircraft Company, McDonnell Douglas Corporation, 7-17 May 1985.

Fig. 16—Projected C-17 schedule

failures observed during testing to estimate what reliability would have been demonstrated had there been time to incorporate planned fixes. This could put compliance measurement on a more objective footing and enhance the credibility of those measurements in the acquisition community. More time for compliance measurements can also extend a contractor's commitment to improve system-level R&M deeper into the acquisition life cycle, past the production decision milestone and IOC.

Extending compliance measurements into operations is one means for increasing the calendar time for measurements without necessarily lengthening the acquisition cycle, a particular concern today. Adoption of new development strategies for critical subsystems (discussed in Sec. V) can also facilitate the start of compliance testing earlier in weapon system DT&E. These two approaches may permit compliance measurements to be made over periods measured in years (one to two or more) rather than several months.

Making compliance measurements over a larger time base should increase statistical significance and consequently the confidence in performance demonstrated during compliance testing.

Making more compliance measurements in operationally oriented environments should stimulate contractors to produce designs more suited to those environments and provide a much better gauge of a system's adequacy because of the closer relationship of the compliance testing environment and the operational environment.

Measuring compliance of production equipment to system-level specifications should remove some of the risk from trying to extrapolate compliance test results of immature DT&E test assets to production articles that may be manufactured differently and that may incorporate many design fixes.

MOTIVATING THE CONTRACTOR

Recognition that compliance will be measured by rigorous testing is a motivating factor for meeting specifications in itself, but another critical aspect is the extent of the contractor's obligation if testing reveals that specifications are not met. With the exception of nine avionics LRUs that were under warranty, the F-16 contractor was not automatically required to fix items not meeting R&M specifications. The fixed-price (incentive) contract target cost had been exceeded by the end of flight testing, and the contractor and the Air Force were sharing additional costs. In such situations, the Air Force program office decides what fixes, R&M or otherwise, to implement with the resources available.

Today, there is a clear trend toward increasing the obligation of contractors for meeting contractual specifications. The contractual commitment for satisfying R&M and availability specifications in the C-17 program, which features a system-level warranty, is an example. The contractor must demonstrate R&M and availability performance at or above thresholds described by growth curves through IOC plus 6 months. Under terms of the warranty, if specifications are not met, the contractor is obligated to initiate corrective action at no additional expense to the government to bring the system into compliance. This means that in principle, at least, the contractor faces unlimited liability if he cannot meet threshold specifications. How this arrangement will be implemented in practice as the program unfolds remains to be seen, but clearly the contractor has a strong motivation to meet R&M and availability thresholds. Moreover, he can earn as much as an additional $12 million, depending on the extent to which thresholds are exceeded during a special, intensive, one-month Operational Readiness Evaluation following FOT&E.

IMPLICATIONS OF NEW CONTRACTING APPROACHES

The greater accountability of contractors for suitability-related characteristics because of recent innovations in specification measures, compliance test design, and contractual instruments for motivating compliance, exemplified by the C-17 contracting arrangement, has significant implications for the Air Force and its contractors in terms of the distribution of risks and responsibilities. The contractor is exposed to more risk because of fixed-price contracts (in the case of the C-17) for full-scale development, ground support equipment, and spares, as well as the greater liability associated with not satisfying R&M and availability specifications under terms of the warranty.

Contractor compliance is more dependent on Air Force actions. Broader specification measures, especially those including induced and no-defect maintenance actions, and compliance measurements in operational settings, are more dependent on the training, manning, and maintenance procedures that Air Force personnel use during the compliance measurement period than in those contracting arrangements that measure only inherent reliability predominantly in a test base setting.

In return for the greater risks assumed by the contractor, the Air Force, in the C-17 program, has ceded certain design responsibilities to the contractor, such as the allocation of subsystem R&M specifications (i.e., deciding the stringency of reliability specifications for individual subsystems and elements of those subsystems) and the selection of ground support equipment, to give the contractor the operating flexibility to satisfy the more comprehensive specifications.[31] There are perhaps increased risks, or at least challenges, associated with the Air Force making more system-level compliance measurements in operational settings away from the controlled environment characteristic of DT&E at a test base. There may be a potentially greater role for the operational test community in contributing to the planning and

---

[31]The coupling of increased accountability with more design authority, characteristic of the C-17 contracting arrangement, is by no means universal, however. In several other programs, we found items under warranty for which the producer of the item has considerable accountability, but in which he must at times depend on the good will of the Air Force and other system contractors to help in solving problems for which he is accountable, but the solutions to which depend on actions that cut across item interfaces.

conduct of such compliance testing. The Air Force also faces additional responsibilities when it must administer system-level warranties.

More extensive contracting arrangements for specifying, testing, and motivating R&M compliance could involve increased front-end program costs to the Air Force. These costs must be balanced against potential operational capability benefits and downstream fiscal benefits from fielding systems having better R&M.[32] A willingness to commit front-end program funds to implement more ambitious contracting approaches for achieving better R&M will be an important litmus test of the Air Force's commitment to R&M.

---

[32]RAND research in progress is analyzing the benefits from improved R&M.

V. USING T&E TO REINFORCE MORE DEMANDING REQUIREMENTS

Testing systems is a primary means for acquiring information that can resolve uncertainties and reduce risks during development and early production. That information is used to support the engineering design process, to support decisionmaking by Air Force and DoD leadership, and to support operational planning for a system's entry into the inventory.[1]

DoD Directive 5000.3, which expresses DoD test policy, states that the objective of operational test and evaluation (OT&E) is to estimate operational effectiveness and suitability, and to provide information on tactics, doctrine, organization, and personnel requirements. Development test and evaluation (DT&E) is to assist the engineering design and development process and to verify the attainment of technical performance specifications and objectives. DT&E and OT&E are often combined to save time and money. Combined testing can, however, include separate DT&E or OT&E test events when needed to satisfy particular test objectives, and always includes separate evaluation and reporting of results.

Improving T&E's contribution to the delivery of operationally suitable systems requires consideration not only of test activities themselves, but also of acquisition policies and procedures, external to the process, that can strongly influence the conduct of testing and the utility of its products. This section discusses how the quality of operational suitability assessments changes through various test phases, as shaped by the acquisition process, identifies alterations in acquisition procedures—both for the development of critical subsystems and for management of the transition from development to production—that can enhance the contribution of testing, and assesses the kinds of changes in testing needed to demonstrate new operating concepts and capabilities.

Some of the options for enhancing the contribution of testing to operational suitability could involve increases in development time and cost. The operational suitability benefits these options offer will have to be measured on a program-by-program basis against the possible additional costs associated with their use.

[1]For a good historical overview of T&E's role in supporting defense decisionmaking, see Booz, Allen, & Hamilton Inc. (1981).

SUITABILITY TESTING

Testing Within Current Acquisition Framework

Although the circumstances of each acquisition program are unique, pressures are often present to get systems of higher advertised capability into the field quickly and to avoid apparently costly gaps in transitioning from development to high-rate production.[2] Program scheduling to achieve these objectives often makes it difficult for the T&E community to make early meaningful inputs to decisionmakers, based on field testing, about a system's operational suitability, and also limits the effective use of test data to influence the operational suitability of systems initially deployed in the field.

Schedule constraints work particularly to the disadvantage of reliability testing because it is inherently different from functional performance testing. Because the reliability of a system is to a large extent the product of stochastic failure events, reliability testing must involve repetitive testing of equipment over time to accumulate enough information to estimate reliability. In contrast, functional performance demonstration is more deterministic—i.e., once an aircraft has demonstrated that it can fly Mach 2, there is no need to make repetitive measurements of that demonstrated performance, although prudence in testing does dictate a methodical, sequential approach to probe the limits of functional performance.

The original F-16 development experience illustrates how events shape the time available for testing, how some subsystem development approaches can influence both the value of early testing and the maturity of the fielded system, and how rapid transitions from development to production limit the productive use of test information.

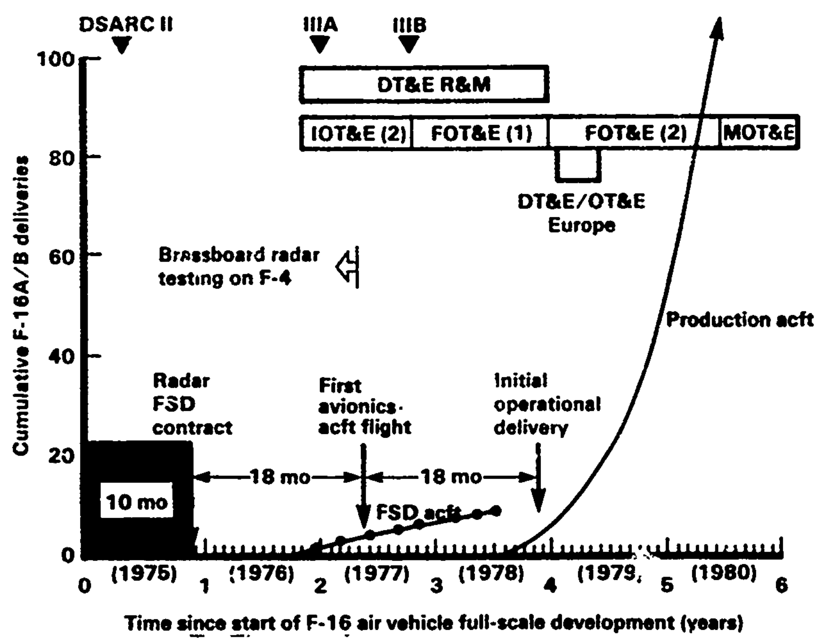
The F-16 program evolved from what was originally a technology demonstration program for lightweight fighter prototypes. Less than half way through that program, the Office of the Secretary of Defense decided to redirect the effort to a competition for the full-scale development of an air combat fighter to satisfy the needs of both the United States Air Force and potential European customers. Willingness to meet early European delivery requirements was a key factor in securing the sale of the aircraft to European customers, but it also resulted in considerable concurrency in development and production activities and meant that many aircraft were produced and delivered before testing was completed.[3]

---

[2]Rich and Dews (1986).

[3]Rich, Stanley, Birkler, and Hesse (1981); and Smith, Barbour, McNaugher, Rich, and Stanley (1981).

The rapid pace of development and deployment in the F-16 program, illustrated in Fig. 17, severely limited opportunities for quantitatively assessing operational suitability before major production milestone decisions. Of particular note is the phasing of the fire control radar subsystem development, one of the most complex and combat critical subsystems on the airplane, which began 10 months *after* the start of air vehicle full-scale development, although there was some flight testing of a prototype (breadboard) air-to-air version of the radar in a competitive flyoff using F-4s prior to that time. Just 18 months after the start of radar full-scale development, a version of the radar was flying in a developmental (pre-production) F-16 aircraft, and just 18 months later production aircraft were being deployed in the field. Because of the phasing of radar development, operational testers never had a mature radar product or associated support elements to test even as late as Follow-on OT&E. The phasing of radar development and the rapid transition to production needed to meet delivery schedules precluded incorporating, before deployment, many fixes to deficiencies identified during testing.



SOURCE: Rich et al. (1981).

Fig. 17—Phasing of equipment development constrains testing and system maturation

In terms of raw flying hours, the F-16 program accumulated considerably more flying hours prior to its initial production decision than did either the A-10 or F-15 (see Table 12), but the kind of equipment undergoing test is also relevant.

Before the initial production decision, prototypes demonstrated the flight-performance potential of the basic air vehicle configuration. These prototypes used many subsystems that were not directly representative of those to be used on the pre-production and production vehicles to come, and lacked some important subsystems such as mission avionics. As a result, assessments could not directly measure the aircraft's operational effectiveness as a weapon system nor its operational suitability, except in a qualitative and projective sense.[4] Just thirteen flight hours were accumulated by the first pre-production airplane prior to the initial production decision, and that vehicle did not incorporate mission avionics. The first development test aircraft with mission avionics accumulated about 143 flying hours prior to the month of the high-rate production decision.

The F-16 program operated under particularly demanding schedules that constrained the assessment of operational suitability during testing, but other Air Force fighter developments have exhibited similar characteristics, suggesting that the F-16 experience was not unique. Figure 18 illustrates that F-15 radar subsystem development began 9

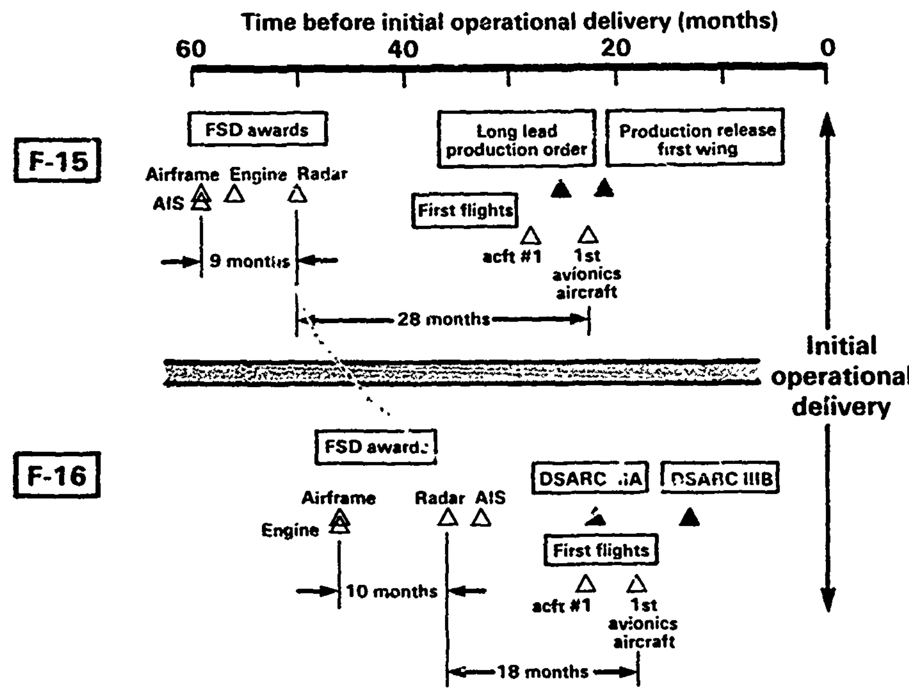
Table 12

TESTING AT MILESTONE III

	Approximate Flight-Test Hours at:				
	Milestone IIIA				
Program	Prototype	FSD	Production	Total	Milestone IIIB
A-10	700	0	0	700	2200
F-15	0	100	0	100	275
F-16	987	13	0	1000	1400
C-17	0	380-450	1200-2200	1600-2700	1600-2700

SOURCE: Rich, Stanley, Birkler, and Hesse (1981); *C-17 Preliminary Design Review (PDR) Reliability and Maintainability*, Douglas Aircraft Company, McDonnell Douglas Corporation, 7-17 May 1985.

[4]This is not meant to denigrate the value of the YF-16 prototypes. They contributed greatly to resolving uncertainties about the feasibility of building a lightweight fighter that possessed desirable air combat characteristics. However, their value for operational suitability assessments was quite limited. See Smith et al., R-2345-AF.



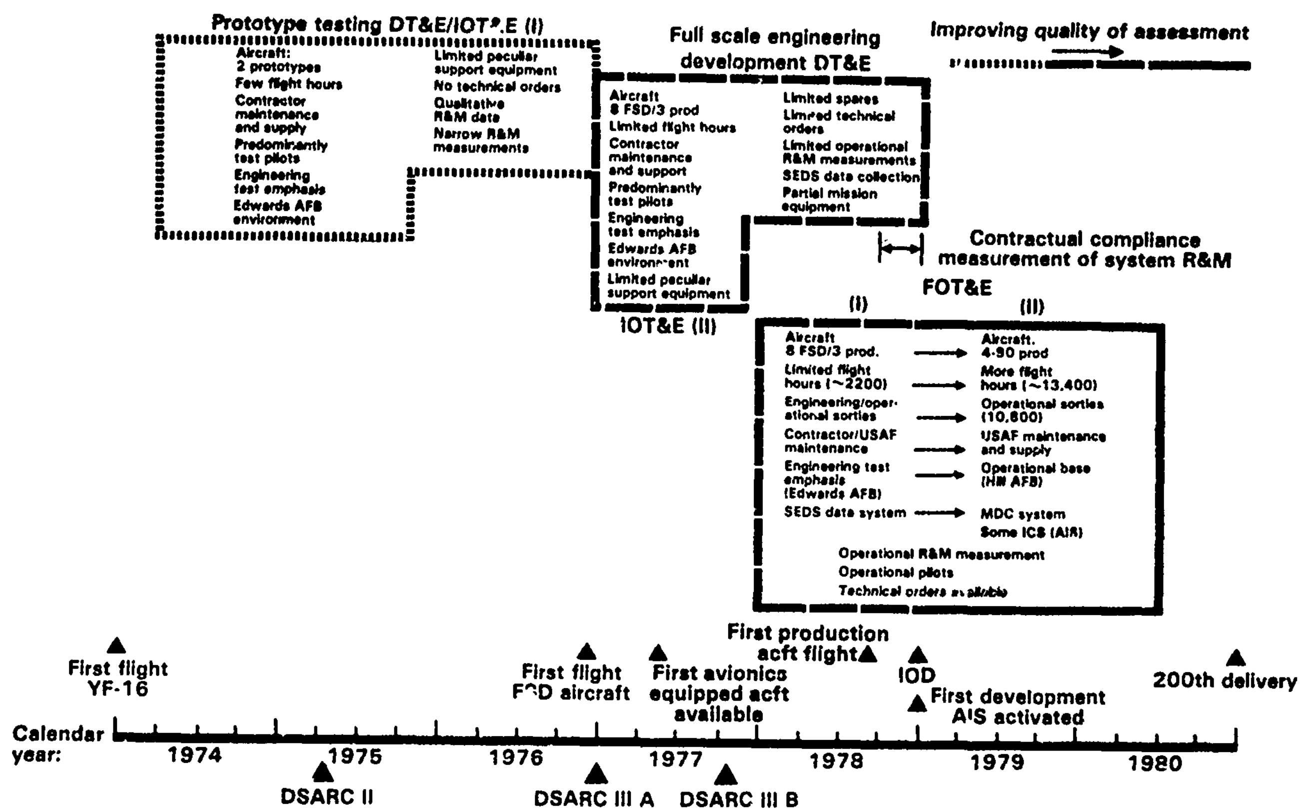
SOURCE: Rich et al. (1981).

Fig. 18—Timing of radar development activities

months after the start of air vehicle development; the initial production decision occurred prior to flight of an aircraft with mission avionics; and the high-rate production decision occurred soon after flight of the first aircraft with mission avionics.

To identify factors that limit the quality of operational suitability assessments through time, we reviewed the progression of operational suitability related testing in the F-16 program (see Fig. 19). Key factors that limited F-16 suitability assessments—and, according to AFOTEC,[5] many other programs as well—included immature or unavailable technical data and support equipment to support the maintenance process, immature software and diagnostic procedures, poor equipment reliability, limited test articles, and nonrepresentative environments for test. In reviewing F-16 testing, we observed that

[5] *Operational Suitability Test and Evaluation*, briefing by Air Force Operational Test and Evaluation Center, undated.

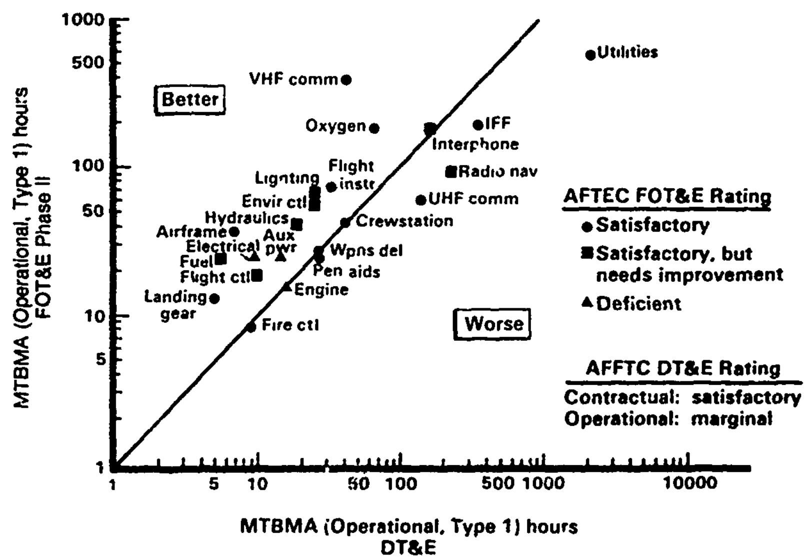


SOURCE: Various F-16 program prototype, DT&E, IOT&E, and FOT&E test reports; Rich et al. (1981).

Fig. 19—Operational suitability assessment through time: F-16A/B program

many of these and other limiting factors only begin to diminish toward the latter phases of Follow-on OT&E (FOT&E), long after production decisions had been made and appreciable numbers of airplanes had been deployed. Hence, the constrained acquisition environment within which testing had to be accomplished severely limited the qualitative and quantitative meaning of suitability testing to support the production decisions.

The immaturity of initially fielded equipment—at least, immaturity relative to perceived levels of satisfactory operational suitability—is apparent on examination of the results of F-16 Follow-on OT&E Phase II, the last stage of F-16 operational testing (see Fig. 20). In comparing F-16 FOT&E Phase II and DT&E test results for the comparatively restrictive reliability measure Mean Time Between Maintenance Actions (Type 1 failures only), essentially an operational measure of inherent reliability, we found that although 55 percent of the subsystems did exhibit improvement between DT&E and FOT&E, an appreciable number of key subsystems showed no significant improvement in reliability. AFTEC rated the reliability of about 45 percent of the subsystems as deficient or only conditionally satisfactory, meaning there



SOURCE: DiGiovanna and Eischens (1979); Wright and Utermohlen (1981); Wright, Utermohlen, and Vraa (1980).

Fig. 20—Systems are still immature late in operational testing

was a significant margin for improvement. Moreover, the absolute levels of reliability were quite low for many subsystems, especially considering that the definition of the reliability measure was very narrow, and that, as we noted in Sec. III, the narrow definition tends to understate the maintenance burden experienced in the field.[6]

A review of the status of deficiency correction efforts provides more insights about the maturity of the F-16 at the conclusion of FOT&E Phase II. Table 13 shows the status of all the service reports that were used to report aircraft deficiencies and were submitted during all F-16 development and operational test programs (1003 of which were submitted during FOT&E). Note that 40 percent of all service reports

Table 13

STATUS OF F-16 SERVICE REPORTS AT THE END OF FOT&E PHASE II

Classification	Category I	Category II	Total
Open	68	825	893
Administratively Closed, Verification Pending	3	43	46
Administratively Closed Verification Pending T.O. Change	4	16	20
Closed	43	1233	1276
Total	118	2117	2235

SOURCE: Wright, Utermohlen, and Vraa (1980).

NOTE: *Category I* service reports are those which describe an emergency condition which presents or has a clear potential to present an unacceptable safety, operational, or maintenance hazard. *Category II* service reports cover everything else. *Open* means corrective action is not resolved and work is continuing. *Administratively Closed, Verification Pending (ACVP)* means the service report is closed but the fix needs to be verified. *Administratively Closed, Verification Pending T.O. Change (ACVT)* means the service report is closed by a technical order change, but the change needs to be verified. *Closed* means corrective action is identified and work is complete with the exception of the change implementation. Closed also includes deficiencies not acted on for reasons of cost or technical feasibility. Service reports are also closed if the ECP approval is complete and verification of the fix is not deemed necessary.

[6]Tracing reliability and other aspects of operational suitability from one phase of testing to another is very difficult because the reliability measures used to report test results differ from one test phase to another. MTBMA (inherent failures only) was the most operationally oriented reliability measure reported in both F-16 DT&E and FOT&E Phase II test reports.

were still open at the end of FOT&E(II), including 58 percent of highest priority in Category I (see note in Table 13 for definitions of terms). Not clear is how many deficiencies were simply accepted and closed out solely for cost or technical reasons. Many design fixes still needed to be implemented at a time when more than 100 aircraft were already in the field and many more were well under way on the production line.[7]

Several factors undoubtedly contribute to the immaturity of fielded equipment, including the late start on the development of key subsystems at the beginning of programs and the pressure to rapidly build up production and deploy systems that tends to outpace deficiency correction efforts. Some of the deficiencies in fielded systems are ultimately corrected, others are not. Recent unpublished RAND research measuring the performance of fighter aircraft radars suggests that many problems tend to linger years after having been identified during operational testing. This suggests that we may have to reassess subsystem development approaches and methods for transitioning from development to production in order to facilitate testing and to make more effective use of its products.

ADJUSTING ACQUISITION STRATEGIES TO ENHANCE THE VALUE OF TESTING[8]

Maturational Development of Critical Avionics

An alternative subsystem development approach can better support early operational suitability assessments by the test community. Currently, contractors often have, from the start of full-scale development, less than two years to develop and install combat-critical avionics in pre-production flight vehicles for testing. Then, in the next two years, they must mature the avionics equipment, demonstrate its functional performance and reliability and maintainability characteristics, and demonstrate the integrated performance and suitability of the

[7]In this program and others, AFTEC and other program participants developed prioritization processess to rank Category II deficiencies (all deficiencies in Category I were by definition high priority). Recent interim guidance changing T.O. 00-35D-54 has institutionalized a number of changes to service report procedures, including procedures for categorizing and prioritizing service reports. See *Interim Guidance T.O. 00-35D-54, Section V*, AF/LEYM, 16 April 1985; *Air Force OT&E Service Reports*, AFTEC briefing, undated.

[8]Other RAND research referenced throughout this subsection explores the issue of alternative acquisition strategies in detail. The intent here is to describe those strategies briefly and illustrate how they can enhance the value of test and evaluation.

avionics suite as a whole. Growing evidence suggests that developing critical avionics on such a time scale is not practical if contractors are to deliver systems that can operate effectively without large amounts of maintenance and logistics support.[9]

This has led to the suggestion that the Air Force recognize the unique development demands of combat-critical avionics and adjust the phasing of weapon system and avionics development accordingly, in a manner somewhat similar to its development approach for aircraft turbine engines. A strategy termed "maturational development" involves decoupling the development of critical subsystems from full weapon system development, and allowing development of those subsystems to begin before that of the full weapon system. Another critical aspect of the strategy is the use of a planned iterative development approach involving multiple development cycles, with a shift in emphasis to R&M improvement after functional performance has been established. The test community plays an important role in identifying engineering and operational deficiencies needing correction as the iterative development process proceeds.

Maturational development has been successfully applied to a number of space and ballistic missile electronic subsystems, as well as to subsystems for commercial applications, yielding both economic and effectiveness benefits. For example, a second development cycle for the Minuteman I inertial navigation system increased its reliability by a factor of 15 and the availability of the missile force rose from 70 percent to over 90 percent. That cycle cost about $150 million (current dollars) and yielded savings of about $1.5 billion.[10] Similar maturational development of critical avionics subsystems for tactical aircraft has rarely occurred, despite the need for improved R&M and fault isolation characteristics for such subsystems.

Greater use of such an approach could (1) provide the test community with a more mature product for use during system-level testing, and (2) present more opportunities for testing associated support elements such as support equipment, technical data, software, and diagnostics. The net result could be the generation of test information that better supports the needs of acquisition decisionmakers.

Despite its apparent advantages, greater adoption of maturational development would force the acquisition community to make even more difficult and complex tradeoffs among functional performance, schedule, cost, and maturity. Its use could add to development time

[9]For a description of equipment characteristics that are a product of this development approach, see Rich and Dews (1986).

[10]RAND research in progress is analyzing other examples of its application.

and front-end program costs, as well as involve tradeoffs between better R&M and functional performance. In the Minuteman I program, the operational capability and economic benefits over the long run were judged to be worth the front-end investment in time and money associated with the maturational development strategy. For more widespread adoption of the strategy, the Air Force will also have to address issues that transcend individual programs, including financial and bureaucratic considerations, as well as Congressional procedures and practices.[11]

Improving the Transition From Development to Production

Virtually all Air Force fighter/attack aircraft programs over the past several decades have moved rapidly from development to high-rate production.[12] There are several reasons this occurs so often. First, if existing systems are unable to meet a perceived threat, the Air Force may be willing to accept an immature product as a stopgap. Second, costs of carrying contractor personnel and facilities can accumulate irrespective of production rates, so there is a natural desire to minimize the time over which costs accumulate. Third, many program managers believe th., the longer the development period, the greater the risk of losing Congressional or DoD support. Fourth, there may be foreign customers for the weapon system if it can be delivered early (this was certainly true in the F-16 program).[13]

The test process suffers under such conditions, of course. When software deficiency correction cycles can take 18 months or more, hardware corrections even longer, and manufacturing lead times may run to 24 months, it is difficult to test amd get results soon enough to improve weapon systems as they first enter the field. Instead, expensive retrofitting is common. For example, work in progress at RAND indicates that radar deficiency corrections for fighter aircraft alone can cost several hundred million dollars. Because of such costs, users sometimes prefer merely to live with deficiencies.

Failure to exploit available test information undermines one of the primary purposes of testing, as well as possible benefits derived from initiatives designed to improve the conduct of testing itself. This has led to suggestions for an alternative acquisition strategy for managing

---

[11]Current RAND research is evaluating the advantages, disadvantages, and implications of maturational development. For an earlier exposition on the subject, see McIver (1974).

[12]See Rich and Dews (1986).

[13]See Rich et al. (1981).

the transition from development to production—a strategy that exploits the knowledge gained from testing before transitioning to high-rate production. "Phased acquisition" delays the start of high-rate production, but not initial low-rate production, by about 18 months over more traditional approaches.[14] Fewer aircraft reach the field during the early years of production, but when aircraft from high-rate production begin to reach the field, they possess more combat capability than they would otherwise. Additionally, because fewer aircraft are fielded during initial low-rate production, fewer must undergo retrofits. A study of the A-7D attack aircraft program estimated that adopting a phased acquisition approach—instead of the more conventional acquisition approach used—might have doubled the sortie-generation capability of the production aircraft over its actual capability. Figure 21 illustrates the essential elements of its hypothetical application to that program.[15]

As advantageous as phased acquisition may be for improving the potential contribution of T&E to the system acquisition process, those benefits must be measured against the factors already mentioned that create pressures for a rapid transition to high-rate production. Several traditional biases must be overcome before comparing phased acquisition and more traditional strategies. First, operators must begin to think about measuring the attractiveness of acquisition strategies and system capability in terms other than "rubber on the ramp" and instead assess the aircraft force's sortie-generation potential and estimated mission success rate. Analysis suggests that the phased procurement approach will lag in sortie-generation potential initially, but then could rapidly overtake the sortie-generation potential delivered by the more traditional acquisition approach as more capable aircraft that benefitted from additional testing are delivered.[16]

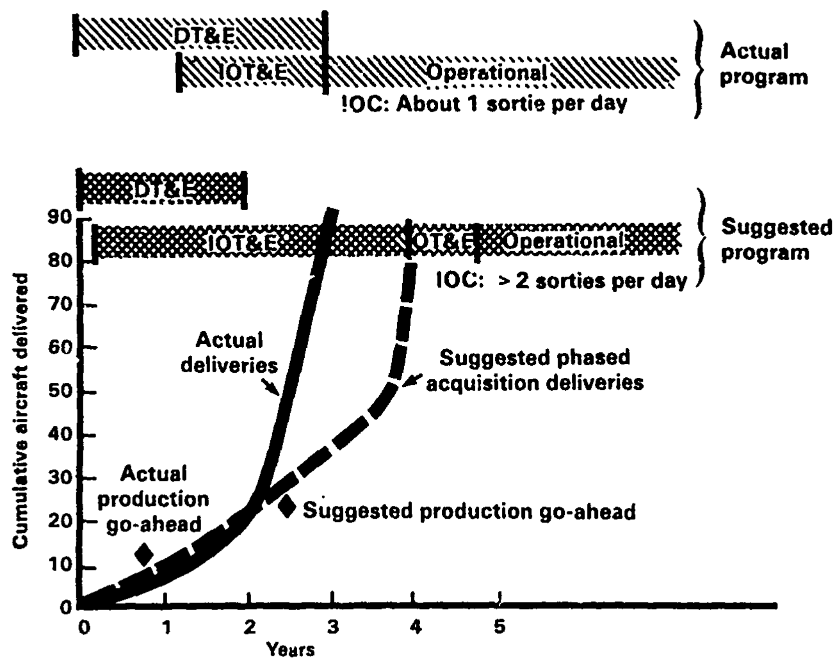
Second, a broader cost perspective must be adopted in comparing phased acquisition with more conventional acquisition approaches. The potentially higher costs from delaying the start of high-rate production must be balanced against the benefits of incorporating fixes during production rather than by retrofit, and against potential operating cost benefits from having a more supportable system in the field.

---

[14]The EF-111A tactical electronic warfare aircraft program included an unplanned phased production program directed by the DSARC at its Milestone III review of the system's readiness for high-rate production. The DSARC, citing unresolved operational suitability problems, ordered additional T&E in parallel with low-rate production to demonstrate the effectiveness of fixes before approving high-rate production. See Booz, Allen, & Hamilton Inc. (1981).

[15]For details on phased acquisition, see Nelson et al. (1974); and Lee (1983).

[16]See Nelson et al. (1974), and Lee (1983) for details.



SOURCE: Adapted from Nelson et al. (1974).

Fig. 21—Applying phased acquisition to the A-7D program

INFLUENCE OF NEW OPERATING CONCEPTS AND CAPABILITIES ON T&E REQUIREMENTS

Future tactical aircraft (such as the Advanced Tactical Fighter) and their associated systems will, for at least part of their life cycle, be operated differently from aircraft of today and will possess new capabilities to support those different operating concepts. Demonstrating that new equipment possesses these new capabilities will entail quantitative and qualitative changes in test and evaluation activities.

There is a general consensus that future systems must possess better reliability and fault isolation characteristics to function effectively in the more stringent operating environments of the future. Improved R&M is a basic policy goal of the Air Force's R&M 2000 program, and reliability has been an important factor in several recent Air Force program decisions.[17] To demonstrate the achievement of these higher

[17] "Air Force Delays Programs to Enforce System Reliability," *Aviation Week & Space Technology*, 9 December 1985, p. 16.

levels of reliability will require the accumulation of more test hours if there is no relaxation in the levels of statistical confidence demanded for test results. Required test hours roughly increase by a factor that is the ratio of the higher reliability to the lower reliability. For systems such as air-to-air missiles, this will necessitate more captive carry testing on aircraft and/or more testing in environmental chambers designed to provide stimuli similar to that experienced in flight. In any event, the quantitative amount of testing will increase, and there is the prospect of increased costs to support increases in testing. Whether this greater amount of testing will also extend development time will depend to some extent on the availability of additional test assets that could permit more rapid accumulation of test data.[18]

New kinds of tests and changes in test emphasis will also be required to demonstrate new operating concepts and capabilities. Table 14 shows some of the more important operating concepts and capabilities and corresponding aspects of testing needing more emphasis. Some of the testing noted in Table 14 is similar to that conducted today, but will be accomplished in new environments. Other tests will be new. Modeling will play a larger role in evaluation efforts than it has in the past, particularly to estimate mobility, attack resilience, and sortie-generation capability.

If the basing flexibility of future systems is to be demonstrated, and particularly the ability to operate in more austere environments, then more testing outside main-operating-base (MOB) environments will be necessary. Most T&E activities are traditionally accomplished at MOBs, first at test bases such as Edwards Air Force Base and later at operating command bases. Because most Air Force airplanes have been designed to operate mainly from MOBs, there has been little need to test in other environments. The only recent exception has been the A-10, which underwent limited testing in simulated Forward Operating Location (FOL) austere environments during IOT&E Phase II and FOT&E Phase I. In the first test, 17 of 713 sorties involved operations at a simulated FOL at Edwards AFB.[19] This test provided a qualitative indication of the A-10's potential when operating from an FOL, but was too short to gauge support needs accurately.

Testing during FOT&E featured a more realistic environment, Gila Bend Air Force Auxiliary Field, and more extended operations, but still

[18] One of the Carlucci defense acquisition initiatives put forward at the beginning of the Reagan administration proposed more front-end funding for test assets to decrease the length of the acquisition cycle. Implementation of this has proven difficult, because of the intense competition for front-end program funding. None of the program offices we visited could point to increases in their allocation of test assets because of the initiative.

[19] Hupp and Mitchell (1976).

Table 14

T&E NEEDING MORE EMPHASIS IN THE FUTURE

Operating Concept or Capability
Operations away from MOBs
Smaller units of aircraft
Short/rough field takeoffs and landings
Less of facilities and support
Streamlined, generalist workforces
Greater mobility and responsiveness
Intertheater deployment
Intratheater dispersal
High, sustained, self-sufficient sortie generation
Resilience to attack
Conventional
Nuclear, chemical, biological
Field maintenance of unconventional materials and structures
Composites
Low observable structures
Aspects of T&E Needing More Emphasis
T&E outside MOB environment
Structured operational field evaluations
Measurement of deployed resource needs
Preparation, loading, unloading, setup-time demonstrations
Loading, deployment modeling
Battle-damage repair demonstrations
Airbase simulation modeling as an integrating tool
Demonstrations of routine maintenance of unconventional materials and structures

far short of the duration expected of future systems. Two aircraft staged a total of 22 sorties (spread across nine different days) through the FOL. A one-day intensive sortie surge demonstration was also conducted at Gila Bend, with two aircraft generating 34 sorties during daylight hours with all flights originating and ending at the FOL.[20]

Although short, these demonstrations were impressive. However, the Air force expects that its future more complex aircraft will need to be able to sustain operations with minimal support at austere sites for weeks at a time. Measuring the minimum levels of manpower, support

[20]Hopkins, Rider, and Mullikin (1977).

equipment, and spares needed to operate in such a manner will require tests of longer duration to more fully exercise the support system and demonstrate basing flexibility.

More structured operational field evaluations, similar to that planned for the C-17 program, will be needed to ensure that systems and their support elements can sustain high sortie-generation rates.[21] This kind of testing will necessarily occur later rather than earlier in operational testing because it requires the availability of a reasonable number of aircraft and support assets. Test ground rules should rigorously prescribe the operating conditions and assets to be used in making the measurements. This implies specifying a time period and base(s) for the demonstration, a specific flying program designed to encompass the spectrum of peacetime and wartime operational missions, and specified levels of manpower, support equipment, and spares support, including specific ground rules for sources of spares and the imposition of supply delays to simulate those that might typically be experienced with the normal logistics system in field operations. The C-17 program's experience with its Operational Readiness Evaluation should be monitored to identify potential pitfalls in implementing such evaluations.

More test measurements will also be needed if mobility is to be evaluated. Some operational test reports we reviewed provided incomplete listings of support equipment requirements for operations in various settings used for testing, but none provided comprehensive listings of equipment needs, the transportation resource requirements to move that equipment, nor the time to prepare, load, unload, set up, and commence operations with that equipment. Providing more comprehensive listings of support equipment needs would appear to be a bookkeeping problem. Determining transportation resource requirements and time requirements will probably require new testing, supplemented with models that integrate the test information so as to make overall assessments of mobility characteristics.

The ability to repair battle damage rapidly in the field is an important feature of a system's resilience to attack and an important factor in sustaining high sortie generation rates, particularly during the early days of a war, before attrition filler aircraft arrive. The ability to accomplish such repairs in the field is expected to become more of an issue in the future as airplane materials and structures change. Consequently, battle-damage repair demonstrations should be considered for inclusion in future testing.

---

[21]For details of the C-17 Operational Readiness Evaluation, see Douglas Aircraft Company (1983).

To test the ability to make routine or battle damage repairs of aircraft that make extensive use of unconventional materials and structures may require the development of new test procedures, because the vector of performance for such aircraft may include many more dimensions, including perhaps the maintenance of low observables.

The operational test community will have to add to its arsenal of models to effectively evaluate the operational suitability of future systems. Today, primary models include the Logistics Composite Model (LCOM) for aircraft systems, an availability analysis model for non-aircraft systems, and mission completion success probability models. We have already noted the need for use of loading and deployment models to evaluate aspects of mobility. Effectively evaluating the resilience to attack of a system and its associated support elements will require use of airbase simulation models that can evaluate how sortie-production changes as various support elements are added or eliminated due to enemy attack or other factors.[22] Developing a capability to use such models to complement field testing requires a long-term commitment on the part of the T&E community and may represent one of the longer lead-time items in Table 14.

ADDRESSING SHORTCOMINGS IN THE TEMP

In examining the T&E process, we identified another aspect of T&E needing attention. The *Test and Evaluation Master Plan (TEMP)* is the primary document used in the OSD review and decision process to assess the adequacy of the testing and evaluation planned for accomplishment during each program phase. It is often regarded as a formal agreement between the parties in the acquisition process about the T&E required during each program phase, and as a means for continuing structured communications about the T&E program among acquisition participants. Previous studies have indicated that it has not been used to maximum effect in past programs.[23] This document will have increasing visibility as the new OSD Office of the Director of Operational Test reviews test plans.

We reviewed *TEMPs* primarily from the perspective of how well they expressed suitability requirements to be tested at various phases in acquisition, and found significant shortcomings (see Fig. 22). These shortcomings are unquestionably created, at least in part, by the chronically poor source requirements documentation we have already

[22]One such model is RAND's TSAR, Theater Simulation of Airbase Resources; see Emerson (1982a).

[23]Booz, Allen, & Hamilton Inc. (1981).

System	Program phase when TEMP issued	Year of TEMP	Sources of reqts noted	MOEs stated	Some explanation of MOEs	Numerical criteria	Goal/ threshold distinction	Interim/ mature distinction
EF-111A	During FSD	1977	No	No	No	No	No	No
TR-1	Pre-FSD	1981	Yes	Few	No	Few	Yes	No
C-X	Pre FSD	1981	Yes	Few	Yes	Few	Yes	Yes
H-X	Pre AFSARC II	1981	Yes	No	No	No	Yes	No
F-16 MSIP	Pre-MSIP test	1981	No	No	No	No	No	No
GBU-15 IR	Pre-FSD	1980	No	No	No	No	No	No
GBU-15 TV	Pre-FOT&E	1983	Yes	Yes	No	Yes	Yes	No
AMRAAM	Pre-D&V	1980/1	Yes	Yes	Yes	Yes	Yes	No
AMRAAM	Pre-FSD	1982	Yes	Yes	Yes	Yes	Yes	Yes

SOURCE: TEMPs as noted.

Fig. 22—Expression of suitability requirements in TEMPS

discussed, and that problem must be addressed to make any headway in improving *TEMP* requirements documentation. One-third of the *TEMPs* we reviewed included no reference for the source of requirements against which system performance during testing would be compared. The program with the best source requirements documentation, the AMRAAM, referenced and attached portions of its *Joint Service Operational Requirement* to the *TEMP* to provide unambiguous documentation of the requirements.

Comparatively few *TEMPs* included the measures of effectiveness (MOEs) to be used during testing, and fewer still defined or explained the meaning of the measures. Comparatively few included the numerical values or criteria for those measures. A significant number of *TEMPs* also failed to distinguish between goals and thresholds and between interim and mature values for requirements. Treatment of requirements in this manner in *TEMPs*, far from facilitating communication among acquisition participants, would seemingly create confusion and uncertainty. The *Baseline Correlation Matrix* discussed earlier may provide a mechanism for improving the treatment of requirements in *TEMPs*, but clearly more discipline in their preparation is desirable.

OBSERVATIONS

For T&E to effectively reinforce more demanding expressions of operational suitability requirements will require changes internal to the T&E process as well as adjustments in traditional acquisition strategy approaches to facilitate testing and to permit the more effective use of T&E products. Internal changes include a need for more testing to demonstrate the achievement of higher levels of reliability, and qualitatively different testing to measure performance away from main operating bases, as well as mobility, sortie generation, and resilience to attack capabilities. An earlier start on the development of combat-critical avionics could help the test community provide decisionmakers with more meaningful early assessments of operational suitability based on field testing. More frequent use of a phased acquisition strategy could facilitate greater use of knowledge gained through testing. Some of these changes, both internal and external to the test process, could require additional time and money, and the Air Force will have to carefully weigh these costs against expected operational suitability benefits derived from their adoption.

VI. FINDINGS AND RECOMMENDATIONS

Developing systems that have the operational suitability characteristics needed to operate successfully in the potentially more stringent and uncertain operating environments of the future will require Air Force initiatives that cut across many functional areas, and include changes in certain aspects of requirements and T&E activities.

Chronic problems in suitability-related requirements expression, and in some cases documentation, begin with the initial expression of operational needs and extend to operational and contractual requirements. These problems are a product of technical, operational, and institutional factors. The action and support of many parts of the Air Force will be needed to ensure consistent improvement in the treatment of operational suitability in requirements expressions.

For T&E to reinforce more demanding expressions of operational suitability requirements will require changes in the conduct of testing and changes in acquisition policies and procedures that shape its conduct and the utility of its products.

The nature and extent of needed operational suitability improvement should be decided on a case-by-case basis by examining the criticality of operational suitability to the military capability being sought. Such examinations can only occur if operational suitability has more visibility in the acquisition process than it has had in the past. The recommendations described below suggest approaches that have the potential for elevating the visibility of operational suitability factors during acquisition. In so doing, they can facilitate consideration of suitability factors in the difficult tradeoff process that must strike the proper balance between operational suitability, functional performance, cost, and development time as acquisition proceeds from the initial expression of needs to operational field testing.

FINDINGS

Historically, SONs have not adequately expressed operational suitability needs, although some improvement is evident in a few recent high-visibility programs. Shortcomings in *SONs*, and particularly in their lack of coverage and quantification of key suitability needs using appropriate measures, have limited their utility to the acquisition community. There are no intractable obstacles to improving the treatment of suitability in future *SONs*, but improvement will require more discipline in the requirements process at all levels from the authors to those ultimately responsible for validating needs. The appreciably better treatment of operational suitability in the *SON* for one-high visibility Air Force program—the Advanced Tactical Fighter—demonstrates that suitability needs can be expressed more effectively if acquisition participants attach enough importance to it.

For formally documenting operational suitability requirements, improved procedures are needed to complement improvements in the substantive expression of requirements. The documentation of operational requirements is fragmented across many sources, requirements are inconsistent from one document to another, and it is extremely difficult to correlate key operational, contractual, and test requirements. The *Baseline Correlation Matrix*, now being applied to selective programs, begins to address some of the basic problems of requirements documentation, but more actions may be needed.

Historically, contractual accountability for suitability-related system characteristics, including R&M, has been quite limited, although there are some encouraging new initiatives. R&M specifications for most types of equipment have typically covered only a fraction of system or subsystem demands for maintenance support. The limited scope and duration of compliance testing, generous test ground rules for excluding failures and maintenance activity, and limited consequences for not meeting specifications have diluted the force of requirements specifications. The C-17 program's contracting approach appears to address many past shortcomings in contractual accountability, although it is too soon to make judgments about its effectiveness.

Acquisition approaches for developing critical subsystems and for managing the transition from development to production limit the contribution of testing. Because the development of combat-critical avionics subsystems has often begun after the start of weapon system full-scale development, initial testing has often been accomplished with key weapon system and support elements either unavailable or immature. This detracts from the quality of operational suitability assessments that support system acquisition decisionmaking.

Rapidly transitioning from development to high-rate production often compromises the effective use of test information to correct deficiencies before fielding systems in substantial numbers. Some deficiencies, residing in aircraft already in the field but initially identified during testing, are corrected by retrofits that are usually more costly than the incorporation of changes on the production line. In other cases, users merely choose to live with degraded capabilities because of the expense of modifying an appreciable number of aircraft already in the field. The failure to exploit available test information undermines the purpose of testing and possible benefits derived from initiatives designed to improve the conduct of testing itself.

More demanding operational suitability requirements will necessitate quantitative and qualitative changes in testing. To demonstrate with confidence the achievement of higher levels of reliability, the amount of testing will have to increase. If new operating concepts and capabilities are to be demonstrated, new kinds of tests, and changes in test emphasis, will be needed.

RECOMMENDATIONS

Improve the treatment of operational suitability in SONs by expanding the aspects of suitability covered, by quantifying key needs using more operationally meaningful measures of military capability outputs desired and significant operating constraints, and by prioritizing key needs. By focusing on the end results required to satisfy a need and the operating constraints within which end results must be achieved, one can identify key suitability characteristics to quantify. This report has illustrated, using a potential fighter aircraft application, how one can then select measures to quantify those characteristics. In prioritizing needs, one should strive to assemble a set of needs that is broad in scope, preferably quantitative, but limited in number so as not to dilute the impact of the prioritization.

Revise policy guidance to explain the rationale for selecting particular measures to define suitability needs and requirements for different types of system applications. Existing guidance identifies and defines many measures, but offers comparatively less assistance for selecting among these many measures a cohesive set that characterizes the major suitability needs for a particular mission application. Guidance is needed that links typical usage patterns of major categories of equipment with appropriate measures. Such guidance could best be incorporated in new Air Force Pamphlet 57-9, *Defining Logistics Requirements in SONs.*

Address the problem of fragmented operational requirements documentation. Institutionalizing the use of the *Baseline Correlation Matrix* document appears desirable to help identify inconsistencies in requirements documentation. Over the long term, development of a unified source of requirements may be desirable that preferably consolidates operational requirements in a single document. With more care in its preparation and review, the existing *System Operational Concept* document might serve this role.

Expand contractual accountability for suitability-related characteristics. This requires (1) using broader R&M specification measures that more fully reflect system and subsystem maintenance demands that drive the Air Force support burden, (2) using compliance-test ground rules that do not compromise those contractual specifications through excessive exclusions of particular kinds of failure events, (3) conducting compliance tests of sufficient duration and scope to have confidence the performance measured is representative of the system at the time of the test, (4) performing more system-level compliance testing on production equipment, (5) conducting compliance tests of system characteristics that contribute to resilience to attack and mobility, and (6) including contractual mechanisms using obligations and incentives to motivate contractors to meet more stringent suitability requirements. The C-17 contracting arrangement embodies many of these requirements and testing initiatives and should be monitored as the program evolves to identify potential implementation pitfalls so that it can be modified and used to support future applications.

Address those aspects of the acquisition process that limit T&E's role in effectively supporting acquisition decisionmaking and in identifying deficiencies and validating their correction. Development of critical subsystems, particularly avionics, should begin before full weapon system development and should use an iterative "maturational development" approach so the test community can have assets of greater maturity to test at the system level before major acquisition decisions. The transition from development to production should be phased such that knowledge gained from testing can be better exploited before transitioning to high-rate production. A "phased acquisition" strategy involves extending the initial low-rate production phase to permit intensive testing and deficiency correction before beginning production at a high rate. Despite the potential of maturational development and phased acquisition for enhancing the contribution of T&E, both face formidable implementation problems involving difficult tradeoffs among system maturity, functional performance, cost, and development time.

Conduct more testing of sortie generation, mobility, and resilience to attack characteristics, and of performance away from main operating base settings to ensure that future systems can operate effectively in more stringent and uncertain operating environments. This will require more structured field evaluations to demonstrate that systems and their support elements can sustain high mission-effective sortie rates. To fully exercise support systems and to demonstrate basing flexibility, tests in austere environments will have to be longer than they have been in the past. Mobility measurements should permit evaluation of responsiveness and transportation resource requirements. Demonstration of the ease of aircraft battle-damage field repair will be necessary as part of any test that measures resilience to attack characteristics. The operational test community will have to develop a capability to use airbase simulation models and loading and deployment models to complement these kinds of tests.

EPILOGUE

This research has identified numerous problems associated with the expression, documentation, and testing of suitability requirements. These problems need attention as part of any cohesive acquisition strategy for improving the operational suitability characteristics of future Air Force weapon systems. Encouragingly, approaches for dealing with many of the problems seem apparent, although their implementation is complicated by the diffuse responsibility for operational suitability within the Air Force, which means that implementing solutions to requirements and testing problems will require actions by many Air Force organizations, rather than a single office.

Air Staff and command organizations set up in recent years as advocates and focal points for R&M (with hopefully a broader view about operational suitability as a whole) can serve an important role by insisting on high standards for the treatment of operational suitability in requirements documents and test plans. Over the long run, this should imbue the requirements and T&E processes with more discipline regarding the treatment of R&M issues and help make R&M a primary consideration in weapon system design, as Air Force leadership has directed.

REFERENCES

"Air Force Delays Programs to Enforce System Reliability," *Aviation Week and Space Technology*, December 9, 1985.

Berman, M. B., *Increasing Future Fighter Weapon System Performance by Integrating Basing, Support, and Air Vehicle Requirements*, The RAND Corporation, N-1985-1-AF, April 1983.

Berman, M. B., *Integrating Basing, Support, and Air Vehicle Requirements: An Approach for Increasing the Effectiveness of Future Fighter Weapon Systems*, The RAND Corporation, R-3304/1-AF, August 1985.

Booz, Allen, & Hamilton Inc., *The Relationship Between Adequacy of Test and Evaluation Data and DSARC Decisions*, June 1, 1981.

Chief of Naval Material, *Operational Availability of Weapon Systems and Equipments: Definitions and Policy*, NAVMAT Instruction 3000.2, 21 January 1981.

Crawford, Gordon, and Cindy Williams, *The Analysis of Subjective Judgment Matrices*, The RAND Corporation, R-2572-1-AF, May 1985.

DiGiovanna, Albert, and W. L. Eischens, *DT&E Reliability and Maintainability Evaluation of the F-16 Aircraft*, Air Force Flight Test Center, AFFTC-TR-79-13, June 1979.

Douglas Aircraft Company, *C-17 Preliminary Design Review (PDR), Reliability and Maintainability*, McDonnell Douglas Corporation, 7–17 May 1985.

Douglas Aircraft Company, *System Specification for C-17 Airlift System*, McDonnell Douglas Corporation, Specification MDC S001, Code Ident 88277, 1 April 1983.

Emerson, D. E., *An Introduction to the TSAR Simulation Program: Model Features and Logic*, The RAND Corporation, R-2584-AF, February 1982a.

Emerson, D. E., *USAFE Airbase Operations in a Wartime Environment*, The RAND Corporation, P-6810, October 1982b.

Emmelhainz, Margaret A., "Innovative Contractual Approaches to Controlling Life Cycle Costs," *Defense Management Journal*, Vol. 19, No. 2, second quarter, 1983, pp. 36–42.

Gabriel, Gen. Charles A., USAF, Chief of Staff, and Verne Orr, Secretary of the Air Force, *Reliability and Maintainability of Air Force Weapon Systems—Action Memorandum*, 17 September 1984.

Hopkins, Maj. Jerry, Maj. William Rider, and Capt. Michael O. Mullikin, *A-10A Follow-on Operational Test and Evaluation: Phase I, Operational Suitability Data Document (Logistics Assessment)*, Air Force Test and Evaluation Center, Report 76-AFTEC-105, May 1977.

Hupp, Col. George C., and Lt. Col. Jay N. Mitchell, *A-10 Initial Operational Test and Evaluation: Phase II Test Report*, Air Force Test and Evaluation Center, Report 76-104, May 1976.

Kern, George, and T. Drnas, *Operational Influences on Reliability*, Rome Air Development Center, RADC-TR-76-366, December 1986.

Lee, A. D., *A Strategy to Improve the Early Production Phase in Air Force Acquisition Programs*, The RAND Corporation, P-6941-RGI, December 1983.

McIver, D. W., *A Proposed Strategy for the Acquisition of Avionics Equipment*, The RAND Corporation, R-1499-AF, December 1974.

Nelson, J. R., et al., *A Weapon System Life-Cycle Overview: The A-7D Experience*, The RAND Corporation, R-1452-PR, October 1974.

Operational Suitability Test and Evaluation, briefing by Air Force Operational Test and Evaluation Center, n.d.

Pyles, Raymond, *The Dyna-METRIC Readiness Assessment Model*, The RAND Corporation, R-2886-AF, July 1984.

Rich, Michael, and Edmund Dews, *Improving the Military Acquisition Process: Lessons from RAND Research*, The RAND Corporation, R-3373-AF/RC, January 1986.

Rich, Michael, W. L. Stanley, and Susan Anderson, *Improving U.S. Air Force Readiness and Sustainability*, The RAND Corporation, R-3113/1-AF, April 1984.

Rich, Michael, W. L. Stanley, J. L. Birkler, and Michael Hesse, *Multinational Coproduction of Military Aerospace Systems*, The RAND Corporation, R-2861-AF, October 1981.

Rue, H. D., and R. O. Lorenz, *Study of the Causes of Unnecessary Removals of Avionics Equipment*, Rome Air Development Center, RADC-TR-83-2, January 1983.

Smith, G. K., A. A. Barbour, T. L. McNaugher, M. D. Rich, and W. L. Stanley, *The Use of Prototypes in Weapon System Development*, The RAND Corporation, R-2345-AF, March 1981.

Wright, Lt. Col. David D., Maj. Victor J. Utermohlen, and Maj. Ronald D. Vraa, *F-16 Follow-on Operational Suitability Test and Evaluation, Phase II Final Report*, Air Force Test and Evaluation Center, AFTEC Report 76-AFTEC-106, October 1980.

Zwanziger, Jack, *Estimates of Reliability During the Test and Evaluation Stage: Some Methodological Considerations*, The RAND Corporation, P-7169, December 1985.